

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 10:17:46 ; Search time 69.28 Seconds
(without alignments)
2100.011 Million cell updates/sec

Title: US-09-819-946-2
Perfect score: 4493
Sequence: 1 MLCYARLVGQLLISCCWA.....NSTEHQASIDYTRRCGST 841

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._viral:*
16: sp._bacteria:*
17: sp._archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	3365.5	74.9	842	11 099PG5	099PG5 mus muscula
2	3361.5	74.8	840	11 0920R8	0920R8 ratu
3	3355.5	74.7	842	11 0925I5	0925I5 mus muscula
4	3348.5	74.5	842	11 0923J9	0923J9 mus muscula
5	3324.5	74.0	842	11 099PG6	099PG6 mus muscula
6	1370	30.5	843	11 0920R7	0920R7 ratu
7	1363	30.3	843	11 0925I4	0925I4 mus muscula
8	1350	30.0	843	11 0923J8	0923J8 mus muscula
9	1170	26.0	858	11 0925D8	0925D8 mus muscula
10	1163	25.9	858	11 0925D8	0925D8 mus muscula
11	1159	25.8	858	11 0925A4	0925A4 mus muscula
12	1147	25.5	858	11 0925D9	0925D9 mus muscula
13	1147	25.5	858	11 091VA4	091VA4 mus muscula
14	1141	25.4	858	11 0923K0	0923K0 mus muscula
15	1060.5	23.6	940	13 073635	073635 fugu rubrip
16	1054	23.5	864	13 073637	073637 fugu rubrip

17	1053	23.4	880	13 073639	073639 fugu rubrip
18	1052.5	23.4	940	13 090WL6	090WL6 spar
19	1021	22.7	848	13 093553	093553 carassius a
20	1014.5	22.6	868	13 073636	073636 fugu rubrip
21	996	22.2	856	13 073638	073638 fugu rubrip
22	975	21.7	877	13 09PW88	09PW88 carassius a
23	949.5	21.1	875	13 073640	073640 fugu rubrip
24	940.5	20.9	844	13 093552	093552 carassius a
25	928.5	20.7	912	11 070410	070410 mus muscula
26	745.5	16.6	855	11 070409	070409 mus muscula
27	737.5	16.4	1156	13 0980C6	0980C6 gallus gall
28	737.5	16.4	1188	13 0980C5	0980C5 gallus gall
29	737.5	16.4	1242	13 0980C4	0980C4 gallus gall
30	732	16.3	1199	11 09EPV6	09EPV6 mus muscula
31	703	15.6	1218	13 090FE3	090FE3 oncorhynch
32	695.5	15.5	983	11 062516	062516 ratu
33	692	15.4	779	11 035269	035269 ratu
34	687.5	15.3	872	4 09H3N6	09H3N6 homo sapien
35	687	15.3	879	11 09QYS2	09QYS2 mus muscula
36	638	14.2	977	13 09PWE1	09PWE1 ictalurus p
37	615.5	13.7	850	11 035189	035189 mus muscula
38	599	13.3	866	11 035268	035268 ratu
39	586.5	13.1	458	11 093555	093555 carassius a
40	586	13.0	803	11 035191	035191 mus muscula
41	561	12.5	852	11 035192	035192 mus muscula
42	547.5	12.2	768	11 035266	035266 ratu
43	542	12.1	695	11 035272	035272 ratu
44	535	11.9	802	11 035271	035271 ratu
45	531.5	11.8	408	13 093558	093558 carassius a

ALIGNMENTS

RESULT	ID	1	PRELIMINARY;	PRT;	842 AA.
099PG5	099PG5				
AC	099PG5				
DT	01-JUN-2001 (TREMBLREL. 17, Created)				
DT	01-JUN-2001 (TREMBLREL. 17, Last sequence update)				
DT	01-DEC-2001 (TREMBLREL. 19, Last annotation update)				
DE	POTATIVE SWEET TASTE RECEPTOR T1R1 (FRAGMENT).				
GN	GPR70.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	11				
RP	SEQUENCE FROM N.A.				
RP	STRAIN=129P3/J;				
RX	MEDLINE=21030739; PubMed=11178737;				
RA	Li X., Inoue M., Reed D.R., Hugue T., Puchalski R.B., Tordoff M.G.,				
RA	Nimolija Y., Beauchamp G.K., Bachmanov A.A.;				
RT	"High-resolution genetic mapping of the saccharin preference locus				
RT	(Sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to				
RT	mouse distal chromosome 4."				
RL	Ensembl: Genes 12:13-16(2001).				
DR	EMBL: AF301162; AAK07022.1; -				
DR	MGD: MGI:1340021; GPR70.				
DR	InterPro: IPR001828; ANF_receptor.				
DR	InterPro: IPR000345; CytC_heme_bind.				
DR	InterPro: IPR000337; GPCR_Mgr.				
DR	Pfam: PF000003; 7tm_3; 1..				
DR	Pfam: PF01094; ANF_receptor; 2.				
DR	PRINTS: PR00248; GPCRMR.				
DR	PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.				
DR	PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 1.				
KW	Receptor.				
FT	NON_TER	842	842		
SO	SEQUENCE	842 AA;	93428 MW;	39739A2FF482D33F	CNC64;

Query Match

74.9%; Score 3365.5; DB 11; Length 842;

Best Local Similarity 73.9%; Pred. No. 5e-275;
Matches 622; Conservative 86; Mismatches 133; Indels 1; Gaps 1;

```

OY 1 MLCTAR-LVGIQLILSCMAFACHSTESSPDFTLPGDVLALGFLPHSGCLOVRRPREV 59
DB 1 MLEMAHLL-LSLQLVYCAVSCQRTSSPGFSLPGDVLALGFLPHSGCLOVRRPREV 60
OY 60 TLCDRSCSFENHGYHLFOAMRLGVEEINNSTALLPNTITGLGYLDVDCSPSANYATLRLV 119
DB 61 TSCDSDSDFNGNGYHLFOAMRFEVEINNSTALLPNTITGLGYLDVDCSSSNVATLRLV 120
OY 120 SLPGHIELOGLLHYSPVLAVIGPDSFNRAATTAALLSPFLVPMISYASSSETLSVK 179
DB 121 AQQGCHLEMOHDLRNHSSKVALIGPDNTDAVTTAALLSPFLVPMISYASSSETLSVK 180
OY 180 RQYPSFLRTIPNDKYOVEMVLLLOKFGMTWISLVGSSSDYQOLGVALENATGOGICI 239
DB 181 RKPSPFLRTIPNDKYOVEMVLLLOKFGMTWISLVGSSSDYQOLGVALENATGOGICI 240
OY 240 AFKDIPEFSAQYDERMOCIMRHLAOCATVVVVFSSROLAFEFESVYLTNLTGKVVWA 299
DB 241 AFKDVVPLSAQYDERMOCIMRHLAOCATVVVVFSSROLAFEFESVYLTNLTGKVVWA 300
OY 300 SEAMALSRHITGVPGIQRIGVNLGVALIQRAVGLKAFEEAVARADKEAPRCHKSGWCS 359
DB 301 SEDMAISTYITVNVPGIIGIGTVLGVALIOQROYGLKEFEESYQAVTGAFCRCPGSGWG 360
OY 360 SNOQLRECOAFEMHMPKIKAFSSSANYAVRAYVAVAHGLHOLLGCSGTCARGVYVW 419
DB 361 TNOQLRECHAFETMMPELGAFSMSNAIVRAYVAVAHGLHOLLGCSGTCARGVYVW 420
OY 420 QLLLEOIHVHFLHKTVAFNDNRDPLSSYNIITAMDNGPKMTFTVLGSGSTSPVOLNTN 479
DB 421 QLLLOQIYKVNFLHKTVAFNDNRDPLSSYNIITAMDNGPKMTFTVLGSGSTSPVOLNTN 480
OY 480 ETKIOMHGKDNVPSKVCSSDCLBHQRYVTFHHCCEFCVCGAGTFLNKSDDLRCQPC 539
DB 481 KKKIOMHGKDNVPSKVCSSDCLBHQRYVTFHHCCEFCVCGAGTFLNKSDDLRCQPC 540
OY 540 GKEEMAPEGSCQCFRTVYVFLREHNSVLAANTLILLLLTGAGLPAHMLDTPVVS 599
DB 541 GKEEMAPEGSCQCFRTVYVFLREHNSVLAANTLILLLLTGAGLPAHMLDTPVVS 600
OY 600 AGRLCFLMLGSLAAGSGSLVGFEPTRPACILROALFALGFTIFLSCLTVRSFOLIT 659
DB 601 AGRLCFLMLGSLAAGSGSLVGFEPTRPACILROALFALGFTIFLSCLTVRSFOLIT 660
OY 660 EKFSTRKVPFFYHAMVONHGAFLFVMSAQLLICTLWLVMTPLPAREYORPPLVME 719
DB 661 EKFSTRKVPFFYHAMVONHGAFLFVMSAQLLICTLWLVMTPLPAREYORPPLVME 720
OY 720 CFEHNSLIGFLIYVNGLLISAFACSYLGDLPENYNAKCVTFSLFNFMSIAFFT 779
DB 721 CFEHNSLIGFLIYVNGLLISAFACSYLGDLPENYNAKCVTFSLFNFMSIAFFT 780
OY 780 ASVYDGKYLPAANMAGSLSSGEGGYFLPKCYVILCRPDINSTEHFQASIODYTRRCG 839
DB 781 ASVYDGKYLPAANMAGSLSSGEGGYFLPKCYVILCRPDINSTEHFQASIODYTRRCG 840
OY 840 ST 841
DB 841 TT 842

```

RESULT 2

O920R8 PRELIMINARY: PRT: 840 AA.

AC O920R8: 01-MAY-1999 (TREMBLrel. 10; Created)
DT 01-MAY-1999 (TREMBLrel. 10; Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19; Last annotation update)
OS Rattus norvegicus (rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR;
RX MEDLINE=99159821; PubMed=10052456;
RA Hoon M.A., Adler E., Lindemeier J., Battey J.F., Ryba N.J.,
RA Zuker C.S.;
RT "putative mammalian taste receptors: a class of taste-specific GPCRs
with distinct topographic selectivity.";
RL Cell 96:541-551(1999).
DR EMBL; AF127389; AAD18069.1;
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 2.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS50259; G_PROTEIN_RECPR_F3_4; 2.
DR Receptor.
FT NON_TER 840
SQ SEQUENCE 840 AA; 93496 MW; 1RCFB7EFC6B845DB CRC64;

Query Match 74.8%; Score 3361.5; DB 11; Length 840;
Best Local Similarity 74.1%; Pred. No. 1.e-274;
Matches 623; Conservative 81; Mismatches 136; Indels 1; Gaps 1;

```

OY 1 MLCTARLVGIQLILSCMAFACHSTESSPDFTLPGDVLALGFLPHSGCLOVRRPREV 60
DB 1 MLEMAHLL-LSLQLVYCAVSCQRTSSPGFSLPGDVLALGFLPHSGCLOVRRPREV 59
OY 61 TLCDRSCSFENHGYHLFOAMRLGVEEINNSTALLPNTITGLGYLDVDCSPSANYATLRLV 120
DB 60 SCDRPSDFNGNGYHLFOAMRFEVEINNSTALLPNTITGLGYLDVDCSSSNVATLRLV 119
OY 121 LFGQHIELOGLLHYSPVLAVIGPDSFNRAATTAALLSPFLVPMISYASSSETLSVK 180
DB 120 LOGPHRIEIOKDLRNHSSKVALIGPDNTDAVTTAALLSPFLVPMISYASSSETLSVK 179
OY 181 QYPSFLRTIPNDKYOVEMVLLLOKFGMTWISLVGSSSDYQOLGVALENATGOGICIA 240
DB 180 KFPSPFLRTIPNDKYOVEMVLLLOKFGMTWISLVGSSSDYQOLGVALENATGOGICIA 239
OY 241 FDIPEFSAQYDERMOCIMRHLAOCATVVVVFSSROLAFEFESVYLTNLTGKVVWA 300
DB 240 FDIPEFSAQYDERMOCIMRHLAOCATVVVVFSSROLAFEFESVYLTNLTGKVVWA 299
OY 301 EAMALSRHITGVPGIQRIGVNLGVALIQRAVGLKAFEEAVARADKEAPRCHKSGWCS 360
DB 300 EDMAISTYITVNVPGIIGIGTVLGVALIOQROYGLKEFEESYQAVTGAFCRCPGSGWG 359
OY 361 NOLRECOAFEMHMPKIKAFSSSANYAVRAYVAVAHGLHOLLGCSGTCARGVYVW 419
DB 360 NOLRECHAFETMMPELGAFSMSNAIVRAYVAVAHGLHOLLGCSGTCARGVYVW 420
OY 421 LEOIHKVHFLHKTVAFNDNRDPLSSYNIITAMDNGPKMTFTVLGSGSTSPVOLNTN 479
DB 420 LEOIHKVHFLHKTVAFNDNRDPLSSYNIITAMDNGPKMTFTVLGSGSTSPVOLNTN 480
OY 481 TKIOMHGKDNVPSKVCSSDCLBHQRYVTFHHCCEFCVCGAGTFLNKSDDLRCQPC 539
DB 480 TKIOMHGKDNVPSKVCSSDCLBHQRYVTFHHCCEFCVCGAGTFLNKSDDLRCQPC 540
OY 541 KEEMAPEGSCQCFRTVYVFLREHNSVLAANTLILLLLTGAGLPAHMLDTPVVS 600
DB 540 KEEMAPEGSCQCFRTVYVFLREHNSVLAANTLILLLLTGAGLPAHMLDTPVVS 599
OY 601 GGRLCFLMLGSLAAGSGSLVGFEPTRPACILROALFALGFTIFLSCLTVRSFOLIT 660
DB 600 GGRLCFLMLGSLAAGSGSLVGFEPTRPACILROALFALGFTIFLSCLTVRSFOLIT 659
OY 661 KFTSKVPFFYHAMVONHGAFLFVMSAQLLICTLWLVMTPLPAREYORPPLVME 720

```



```

||||| 1 |||||||: 1 |||||||: 1 |||||||: 1 |||||||: 1 |||
Db 660 KESTKVPTEFYHTAQNAGAGLFVIVSSVTHLLICITLWLMVMTPTPTREYQCFPHLVILEC 719
Qy 721 TENNSIGFLIAPLYNGLISTSAFACSYLGGDLPENYNEAKCVTFSLTFNPSWIAFPTM 780
Db 720 TENNSVGFLLAFHNLISTFVCSYLGKELPENYNEAKCVTFSLTFNPSWIAFPTM 779
Qy 781 SVYDGKYLPAANMAGLSLSSGFGYFLPKCYIILCRPDNSTEHFQASIQDYTRRCG 840
Db 780 SITVGSFLPANNVLAGLITLSSGFSGYFLPKCYIILCRPELNTNHEHQASIQDYTRRCGT 839
Qy 841 T 841
Db 840 T 840

RESULT 3
Q92515 PRELIMINARY; PRT; 842 AA.
AC Q92515;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CANDIDATE TASTE RECEPTOR T1R1.
GN T1R1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21219400; PubMed=11319557;
RA Montmayeur J.P., Liberles S.D., Matsunami H., Buck L.B.;
RT "A candidate taste receptor gene near a sweet taste locus.";
RL Nat. Neurosci. 4:492-498(2001).
EMBL: AF337040; AAK39437.1; -.
KM Receptor.
SQ SEQUENCE 842 AA; 93471 MW; FDBA0CC1BD45DEE7 CRC64;

Query Match 74.7%; Score 3355.5; DB 11; Length 842;
Best Local Similarity 73.6%; Pred. No. 3.5e-274;
Matches 620; Conservative 87; Mismatches 134; Indels 1; Gaps 1;

Qy 1 MLCSTAR-LVGLQLLISCCNAFACHSTESSPDTFLPGDYLLAGLFPPLHSGCLOYRHPREV 59
Db 1 MLEFAAHLLLSLQAVAYCMAFSCORTSSPFGSLPGDFLLAGFLSHADCLQVRRHRLV 60
Qy 60 TLCDRSCSFNHHGYHLFOAMRLGVEEINNSSTALLPNTITLGOLYDVCSDSANVATLRLV 119
Db 61 TSCDRSDFNHHGYHLFOAMRLGVEEINNSSTALLPNTITLGELDYDVCSSSNVATLRLV 120
Qy 120 SLPGOHIELOGDILHSPVLAIVIGPSTNRATTAALSPFLVPMISYASSETLSVK 179
Db 121 AAGGTGHEMQORDLRNHSKVALIGPDNDHATTAALSPFLMPLVYASSVILSGK 180
Qy 180 ROYPSFLRTIPNDKYOVETWVLLQKFGWTWISLVGSSDDYGOLGVQALENQATGOGICI 239
Db 181 RKPSFLRTIPDSKYOVETVIRLQSGFWWISLVGSGDYGOLGVQALENQATGOGICV 240
Qy 240 AFKQIMPFSAOVDEROQCLMRHLAOGATVVVYSSRQALAFVEESVVLNLTGKWVYA 299
Db 241 AFKDVVLSAOGDPPRQMRMLRLARARTVVVYVFSNRHLAGVFFRSVVLNLTGKWVYA 300
Qy 300 SEANALSRHITGVGIGRIQGVAVLQKRAVPLGAKFEAAYARADKAPRCHKGSWCS 359
Db 301 SEDMAISTYITNVPGIGIGIVLGAIQOROVPGIKKEPEESYVOAVMGAPRTCPGSGWC 360
Qy 360 SNQCRCCQAFMATMTRKIAFSMSAYNAYRAYVAHGHQHLGCASGCSGRYVPM 419
Db 361 TNOICRCHAFETWNNMDELGAFSMSAAYNAYEAVYAAHGHQHLGCTGCTGARGPYVPM 420

```

```

Qy 420 QLLQEIHKVHFLHKDYAFANDNRDPLSSYNIIAMDNMGPKMTFTVLGSSWSPVOLIN 479
Db 421 QLLQOIYKVNFLHKHTYAFNDKDPGLGYDIIAMDNGPKMTFTVIGSASLSYPVHLIN 480
Qy 480 ETKIOMHGKDNQPKSVCCSCLBGNORVYTGFIHHCCEYVPCGAGTFLNKSIDLYRCOPC 539
Db 481 KTKIOMHGKNNQVSVCTRCLEGNHRLVMSGHHCCEPCAGTFLNSELTCQPC 540
Qy 540 GKEEMAPGSGOTCEPRTVFLALREHTSWVLANTLILLLILLAGLFAHMLDPVYRS 599
Db 541 GTEEMAPGSSACFSRTYEFGLGMEHPISLVLAANTLILLLILLAGLFAHMLDPVYRS 600
Qy 600 AGGRICFLMLGSLAAGSGSLYGFGEPTRPACLLRQALFALGFTTFLSCLTVRSQOLI 659
Db 601 AGGRICFLMLGSLAAGSGSLYGFGEPTRPACLLRQALFALGFTTFLSCLTVRSQOLI 660
Qy 660 KESTKVPTEFYHTAQNAGAGLFVIVSSVTHLLICITLWLMVMTPTPTREYQCFPHLVILEC 719
Db 661 KESTKVPTEFYHTAQNAGAGLFVIVSSVTHLLICITLWLMVMTPTPTREYQCFPHLVILEC 720
Qy 720 CTENNSIGFLIAPLYNGLISTSAFACSYLGGDLPENYNEAKCVTFSLTFNPSWIAFPTT 779
Db 721 CTENNSIGFLIAPLYNGLISTSAFACSYLGGDLPENYNEAKCVTFSLTFNPSWIAFPTT 780
Qy 780 ASYVDGKYLPAANMAGLSLSSGFGYFLPKCYIILCRPDNSTEHFQASIQDYTRRCG 839
Db 781 SITVGSFLPANNVLAGLITLSSGFSGYFLPKCYIILCRPELNTNHEHQASIQDYTRRCG 840
Qy 840 ST 841
Db 841 TT 842

```

```

RESULT 4
Q92309 PRELIMINARY; PRT; 842 AA.
AC Q92309;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TASTE RECEPTOR T1R1.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
RA Zuker C.S.;
RT "Mammalian Sweet Taste Receptors.";
RL Cell 10:0-0(2001).
EMBL: AY032622; AAK51603.1; -.
KM Receptor.
SQ SEQUENCE 842 AA; 93455 MW; D4D2CE90959E991A CRC64;

```

```

Query Match 74.5%; Score 3348.5; DB 11; Length 842;
Best Local Similarity 73.5%; Pred. No. 1.3e-273;
Matches 619; Conservative 87; Mismatches 135; Indels 1; Gaps 1;

Qy 1 MLCSTAR-LVGLQLLISCCNAFACHSTESSPDTFLPGDYLLAGLFPPLHSGCLOYRHPREV 59
Db 1 MLEFAAHLLLSLQAVAYCMAFSCORTSSPFGSLPGDFLLAGFLSHADCLQVRRHRLV 60
Qy 60 TLCDRSCSFNHHGYHLFOAMRLGVEEINNSSTALLPNTITLGOLYDVCSDSANVATLRLV 119
Db 61 TSCDRSDFNHHGYHLFOAMRLGVEEINNSSTALLPNTITLGELDYDVCSSSNVATLRLV 120
Qy 120 SLPGOHIELOGDILHSPVLAIVIGPSTNRATTAALSPFLVPMISYASSETLSVK 179
Db 121 AAGGTGHEMQORDLRNHSKVALIGPDNDHATTAALSPFLMPLVYASSVILSGK 180
Qy 180 ROYPSFLRTIPNDKYOVETWVLLQKFGWTWISLVGSSDDYGOLGVQALENQATGOGICI 239

```

```

Db 181 RRFPSFLRTIPSDKYQVEIVYVRLQSFQVWVMSLVGSYDYDQGLVQALDEELATRGICV 240
OY 240 AFKDIIMPESAQVDERMOCIMRLHQAQATVYVVSRLAVFESVYLTNLTKWVWA 299
Db 241 AFKDVVPLSAQAGDPRMOMMLRLARFTVVVVSNNRLAVFSSVYLVLAFLHPVVS 300
OY 300 SEAMALSRHITGVPGIORIGVAVIAOKRAVPGKAFEEAARADKEAPRCHGSMCS 359
Db 301 SEDMAISITVITVPCIGIGTVLGAIOQRQVPGKKEFEESVYQAVMGPRTCPBESWMC 360
OY 360 SNOLORECOAFPAHMPKLIKAFSSSAYNAVAVAHGLHQLIGCASGACSRGRVYW 419
Db 361 TNOLORECHAFPTMMPELGAFSMSAAYNAVAVAHGLHQLIGCTSGTCARGPVYW 420
OY 420 OLLEOIHVHFLHDKDTVAFNDNRPLSSYNIITANDMNGPKFTFYLGSTMSPVOLNIN 479
Db 421 QLLQOIYKVNFLHKKTYAFDDKDPGLGYDIIANDMNGPEMTFEVIGSASLSPVHLIDIN 480
OY 480 ETKIOMHGKDNQVPRKSVCSDDLGEHQRVYVTFHHCCEPCVCGAGTFPLNKSDLYRCOPC 539
Db 481 KTKIOMHGKNNQVPRSVCTRDCLGEHRLVWMSHHCCFECMPCGAGTFPLNTSELHTCOPC 540
OY 540 GKEENAPRESSQCFPRTYVFLALREHTSVVLAANTLLILLGTAGLFAMHLDPVVS 599
Db 541 GKEENAPRESSACFSRTVEFLGMHBPISLVLLAANTLLILLGTAGLFAMHLDPVVS 600
OY 600 AGGRICFLMLGSLAAGSGSLYGFGEPTPACILROALFALFTFPLSCLYRSOLYII 659
Db 601 AGGRICFLMLGSLVAGSGSLYGFGEPTPACILROALFALFTFPLSCLYRSOLYII 660
OY 660 FKFSTRKVPFTYHAWVQNHAGLFWMISSAOQLICLTWLVVTPPLPARYOFRPHLYME 719
Db 661 FKFSTRKVPFTYHAWVQNHAGLFWMISSAOQLICLTWLVVTPPLPARYOFRPHLYME 720
OY 720 CRETSLGFIILAFVINGLISAFACSYGKDLPENYNAKCVTSLSLEFNYSIAFTT 779
Db 721 CRETNSVGLVFAHNILISTIFWCSTYGLKELPENYNAKCVTSLSLEFNYSIAFTT 780
OY 780 ASVYGKTYLPAANMAGLSLSSGFGYFLPRCYVILCRPDNSTEHFOASIODYTRRCG 839
Db 781 SSITGSGYLPANVNLAGLATLSGFGSYFLPRCYVILCRPELNTEHFOASIODYTRRCG 840
OY 840 ST 841
Db 841 TR 842

```

```

DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm.3.1.
DR PRINTS: PR00248; GPCRMR.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS00259; G_PROTEIN_RECIP_F3_4; 1.
DR RECEPTOR.
DR NON_TER.
DR SEQUENCE 842 AA; 93425 MW; 58826C43F5D0352E CRC64;

Query Match 74.0%; Score 3324.5; DB 11; Length 842;
Best Local Similarity 73.0%; Pred. No. 1.4e-271;
Matches 615; Conservative 89; Mismatches 137; Indels 1; Gaps 1;

OY 1 MLCIAR-LVGLQILSCIOAFACSTESSPDFTLPGDVLGLAPRLSSCLOVHRREV 59
Db 1 MLEMAHLHLSLQILAVYCMASFSCORTSSPEFSLEPDLGLGLSLHADCLOVHRREV 60
OY 60 TLCDRSCSFNEGHYLFQAMRLGVEEINNSTALLPNIITLGYOXYDCSDSANVATRLV 119
Db 61 TSCDRSDSPNGHYHLPQAMRFTVEEINNSTALLPNIITLGYELDYDCSSNVATRLV 120
OY 120 SLPGCHHIFLQDGLHSTYVAVIGPDSNRAATTAALSPFLVPMISAASSETLSVK 179
Db 121 AQQGTHLEMOHDLNRHSSKVVYALIGPDMDHAVTAAALSPFLMPLVSEASSVLSGK 180
OY 180 RYPSFLRTIPNDKYQVEIVYVRLQSFQVWVMSLVGSSPDYDQGLVQALDEELATRGICV 239
Db 181 RRFPSFLRTIPSDKYQVEIVYVRLQSFQVWVMSLVGSYDYDQGLVQALDEELATRGICV 240
OY 240 AFKDIIMPESAQVDERMOCIMRLHQAQATVYVVSRLAVFESVYLTNLTKWVWA 299
Db 241 AFKDVVPLSAQAGDPRMOMMLRLARFTVVVVSNNRLAVFSSVYLVLAFLHPVVS 300
OY 300 SEAMALSRHITGVPGIORIGVAVIAOKRAVPGKAFEEAARADKEAPRCHGSMCS 359
Db 301 SEDMAISITVITVPCIGIGTVLGAIOQRQVPGKKEFEESVYQAVMGPRTCPBESWMC 360
OY 360 SNOLORECOAFPAHMPKLIKAFSSSAYNAVAVAHGLHQLIGCASGACSRGRVYW 419
Db 361 TNOLORECHAFPTMMPELGAFSMSAAYNAVAVAHGLHQLIGCTSGTCARGPVYW 420
OY 420 OLLEOIHVHFLHDKDTVAFNDNRPLSSYNIITANDMNGPKFTFYLGSTMSPVOLNIN 479
Db 421 QLLQOIYKVNFLHKKTYAFDDKDPGLGYDIIANDMNGPEMTFEVIGSASLSPVHLIDIN 480
OY 480 ETKIOMHGKDNQVPRKSVCSDDLGEHQRVYVTFHHCCEPCVCGAGTFPLNKSDLYRCOPC 539
Db 481 KTKIOMHGKNNQVPRSVCTRDCLGEHRLVWMSHHCCFECMPCGAGTFPLNTSELHTCOPC 540
OY 540 GKEENAPRESSQCFPRTYVFLALREHTSVVLAANTLLILLGTAGLFAMHLDPVVS 599
Db 541 GKEENAPRESSACFSRTVEFLGMHBPISLVLLAANTLLILLGTAGLFAMHLDPVVS 600
OY 600 AGGRICFLMLGSLAAGSGSLYGFGEPTPACILROALFALFTFPLSCLYRSOLYII 659
Db 601 AGGRICFLMLGSLVAGSGSLYGFGEPTPACILROALFALFTFPLSCLYRSOLYII 660
OY 660 FKFSTRKVPFTYHAWVQNHAGLFWMISSAOQLICLTWLVVTPPLPARYOFRPHLYME 719
Db 661 FKFSTRKVPFTYHAWVQNHAGLFWMISSAOQLICLTWLVVTPPLPARYOFRPHLYME 720
OY 720 CRETSLGFIILAFVINGLISAFACSYGKDLPENYNAKCVTSLSLEFNYSIAFTT 779
Db 721 CRETNSVGLVFAHNILISTIFWCSTYGLKELPENYNAKCVTSLSLEFNYSIAFTT 780
OY 780 ASVYGKTYLPAANMAGLSLSSGFGYFLPRCYVILCRPDNSTEHFOASIODYTRRCG 839
Db 781 SSITGSGYLPANVNLAGLATLSGFGSYFLPRCYVILCRPELNTEHFOASIODYTRRCG 840
OY 840 ST 841

```

Db 841 TT 842

RESULT 6

0920R7 PRELIMINARY: PRT: 843 AA.

AC 0920R7:

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE PUTATIVE TASTE RECEPTOR TR2 (FRAGMENT).

OS Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MISTAR.

RX MEDLINE=99159821; PubMed=10052456;

RA Hoon M.A., Adler E., Lindemeier J., Battey J.F., Ryba N.J., Zuker C.S.,

RT "Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity."

RL Cell 96:541-551(1999).

DR EMBL; AF127330; AAD18070.1; -.

DR InterPro: IPR001828; ANF_Receptor.

DR Pfam; PF00003; 7tm_3; 1.

DR Pfam; PF01094; ANF_Receptor; 2.

DR PRINTS; PR00248; GPCRMR.

DR PROSITE; PS50259; G_PROTEIN_RECEPTOR_F3_4; 2.

KW Receptor.

FT NON_TER

SO SEQUENCE 843 AA; 95799 MW; D23AC22D1E04988 CRC64;

Query Match 30.5%; Score 1370; DB 11; Length 843;
Best Local Similarity 37.8%; Pred. No. 1e-106;
Matches 318; Conservative 145; Mismatches 316; Indels 62; Gaps 22;

Db 841 TT 842

RESULT 6

0920R7 PRELIMINARY: PRT: 843 AA.

AC 0920R7:

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE PUTATIVE TASTE RECEPTOR TR2 (FRAGMENT).

OS Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MISTAR.

RX MEDLINE=99159821; PubMed=10052456;

RA Hoon M.A., Adler E., Lindemeier J., Battey J.F., Ryba N.J., Zuker C.S.,

RT "Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity."

RL Cell 96:541-551(1999).

DR EMBL; AF127330; AAD18070.1; -.

DR InterPro: IPR001828; ANF_Receptor.

DR Pfam; PF00003; 7tm_3; 1.

DR Pfam; PF01094; ANF_Receptor; 2.

DR PRINTS; PR00248; GPCRMR.

DR PROSITE; PS50259; G_PROTEIN_RECEPTOR_F3_4; 2.

KW Receptor.

FT NON_TER

SO SEQUENCE 843 AA; 95799 MW; D23AC22D1E04988 CRC64;

Db 841 R 841

RESULT 7

092514 PRELIMINARY: PRT: 843 AA.

AC 092514:

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE CANDIDATE TASTE RECEPTOR TR2.

GN TR2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J.

RX MEDLINE=21219400; PubMed=11319557;

RA Montmayeur J.P., Liberles S.D., Matsunami H., Buck L.B.;

RT "A candidate taste receptor gene near a sweet taste locus."

RL Nat. Neurosci. 4:492-498(2001).

DR EMBL; AF337041; AAK39438.1; -.

KW Receptor.

SO SEQUENCE 843 AA; 95735 MW; 0543834EAE7AC2E CRC64;

Query Match 30.3%; Score 1363; DB 11; Length 843;
Best Local Similarity 37.8%; Pred. No. 4e-106;
Matches 316; Conservative 142; Mismatches 325; Indels 54; Gaps 19;

Db 841 R 841

RESULT 7

092514 PRELIMINARY: PRT: 843 AA.

AC 092514:

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE CANDIDATE TASTE RECEPTOR TR2.

GN TR2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J.

RX MEDLINE=21219400; PubMed=11319557;

RA Montmayeur J.P., Liberles S.D., Matsunami H., Buck L.B.;

RT "A candidate taste receptor gene near a sweet taste locus."

RL Nat. Neurosci. 4:492-498(2001).

DR EMBL; AF337041; AAK39438.1; -.

KW Receptor.

SO SEQUENCE 843 AA; 95735 MW; 0543834EAE7AC2E CRC64;

OY 777 FTASVYDGKYLPPANMMAGLSSSGFGCYLPKCYVILCRPDLNSTEHF 827
 DB 784 VPLLANVOAYOPAVOMGAILVLCALGILVTFLHPKCYVILMLPKLNTOEFF 834

RESULT 11

OY 925M4 PRELIMINARY; PRT; 858 AA.
 AC Q925M4;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE TASTE RECEPTOR.
 GN TIR3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV;
 RA Sainz E., Korley J.N., Battey J.F., Sullivan S.L.;
 RT "Identification of a novel member of the T1R family of putative taste
 RT receptors."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY026318; AAK01937.1; .
 KW Receptor.
 SQ SEQUENCE 858 AA; 94567 MW; A7879C015C0E8474 CRC64;

Query Match 25.8%; Score 1159; DB 11; Length 858;
 Best Local Similarity 32.9%; Pred. No. 6,6e-89;
 Matches 273; Conservative 150; Mismatches 354; Indels 54; Gaps 16;

OY 29 SPDFLLPGDYLLAGLFPPLHS---GCLQVRHREPVYLCDSKSCFNEGHYLPQAMRLGVEE 85
 DB 26 SOQFKAQGDYILIGLFPPLGSTEATLNTQAPNSTLCNR--FSPGLFLAMAKMAVEE 82
 OY 86 INNSTALLPNTLGYOLYVCSDS-ANYATLRVLSLPGQNHIELOGDLHSPYLAIV 144
 DB 83 INNGSALLPGLRLGDLFTDTCSEPVYTKSSLMFLAKVSOSIAVCNTQYQPRYLAIV 142
 OY 145 GPDSTNRATTAALLSPFLVPMISYAASSETLSVKRQYPSFLRTIPNDKYQYETVNLQ 204
 DB 143 GPHSSELLITLTKFFSFILMPQVYSASMDRLSDRETFPSFRTVPSDRVQLQAVYTLQ 202
 OY 205 KFGWTWISLVGSDDYGGQLGVALENOATGQICIAFKDIMPFSQVQDE--RMQCLMRH 262
 DB 203 NFSMNVAALGSDDDYGRGLSTFSSLANARGLCTAHBGLVFOHDTSGOOLGRVLDVLRQ 262
 OY 263 LAQAGATVVVVFSSRQLARVFESVYVLTNLGKVVVASEAMALSRHITGPGIORIGVL 322
 DB 263 VNOSKYQVVVLFASARAVYSLEFSYSIHGSLSPKVVVASESWLTSDLVMTLPINARGYVL 322
 OY 323 GVAIQRAVPGILKAEEAYARADKEAPRCHGKSCSS-----NOLCRECOAF 370
 DB 323 GFLQGRALLPERSHYETHLAL--AADP---AFCASLNAELDEEHVWGRCRCDI 375
 OY 371 MAHTMPKLKAFSSSA-----YNAYRAVYVAHGLHQLGASGACS-GRGYVPQOLLE 423
 DB 376 MIONSSSGLQLNSAGOLHJOIFATYAAVYSAVALHNTLQCNVSHCHVSEHYLPQOLLE 435
 OY 424 QIHKVFLLHKDVTAFENDNRDLSSYNIITAMDNGPKMTFTVLGSGSTWSPVQNLINETI 483
 DB 426 NMYNNSFHARDLTLQFDAGNVDMEDLKMMVWQSPPTPLVHTVGFNGT---LQLOQSKM 492
 OY 484 OMHGNDNVPKSVCSDDLGEHQRYVTGPHHCCFECVPCGASTFLKSLYLKQCPGAKKE 543
 DB 483 YWPG--NOVPVQSCROCKDQGVRRYKGFHSCCYDVCDAKASYRKHNPDPFCTPCNOQO 550
 OY 544 WAPESQCPPTVYVLLALREHTSWLLAANTLILLGLTGAGLFAWMLDTPVVSAGS- 602
 DB 551 WSPKSTACLPKRPKFLMAGEPVVLSLLLLCTLVGLALALGLSVNHMDSPVLOASGGS 610

OY 603 RLCE--LMLGSLAAGSGSLYGFGEPTRPACLLRQALFALGFTIFLSCVTSFQOLIIF 660
 DB 611 OCFEGLICGLPCL---SVLLFPGRRSSASCLAQCPMAHLPLTGCLSTFLQAAETFEV-- 665
 OY 661 KESTKVPFTFYHAMVONHAGL-----VMISAAQOLICLTWLVVMPPLAREQRFPHLY 716
 DB 666 --ESELPLSMAMWLCGYLGLGAMVLLATVFEALCAMYLTAFPEVVTQWSVLPTEV 723
 OY 717 MLECTNSLGLFIATLVNGLSLISAFACSVYGLKDLPEVYNKACTFSLTFNFSMAIF 776
 DB 724 LEHCYRSMVSLGLVIHTNMLAFLCFLGLFVLQSQPGKRNRRGLTFPMLAFITWVSF 783
 OY 777 FTASVYDGKYLPPANMMAGLSSSGFGCYLPKCYVILCRPDLNSTEHF 827
 DB 784 VPLLANVOAYOPAVOMGAILVLCALGILVTFLHPKCYVILMLPKLNTOEFF 834

RESULT 12

OY 925D9 PRELIMINARY; PRT; 858 AA.
 AC Q925D9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE SWEET TASTE RECEPTOR FAMILY 1 MEMBER 3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=2125287; PubMed=11326277;
 RA Max M., Shanker Y.G., Huang L., Rong M., Campagne F., Liu Z.,
 RA Weinstein H., Damak S., Margolskee R.F.;
 RT "Tas1r3, encoding a new candidate taste receptor, is allelic to the
 RT sweet responsiveness locus Sac.";
 RL Nat. Genet. 28:58-63(2001).
 DR EMBL; AF368024; AAK5536.1; .
 KW Receptor.
 SQ SEQUENCE 858 AA; 94633 MW; 4E416824FDB478F8 CRC64;

Query Match 25.5%; Score 1147; DB 11; Length 858;
 Best Local Similarity 32.6%; Pred. No. 6,8e-88;
 Matches 271; Conservative 150; Mismatches 356; Indels 54; Gaps 16;

OY 29 SPDFLLPGDYLLAGLFPPLHS---GCLQVRHREPVYLCDSKSCFNEGHYLPQAMRLGVEE 85
 DB 26 SOQFKAQGDYILIGLFPPLGSTEATLNTQAPNSTLCNR--FSPGLFLAMAKMAVEE 82
 OY 86 INNSTALLPNTLGYOLYVCSDS-ANYATLRVLSLPGQNHIELOGDLHSPYLAIV 144
 DB 83 INNGSALLPGLRLGDLFTDTCSEPVYTKSSLMFLAKVSOSIAVCNTQYQPRYLAIV 142
 OY 145 GPDSTNRATTAALLSPFLVPMISYAASSETLSVKRQYPSFLRTIPNDKYQYETVNLQ 204
 DB 143 GPHSSELLITLTKFFSFILMPQVYSASMDRLSDRETFPSFRTVPSDRVQLQAVYTLQ 202
 OY 205 KFGWTWISLVGSDDYGGQLGVALENOATGQICIAFKDIMPFSQVQDE--RMQCLMRH 262
 DB 203 NFSMNVAALGSDDDYGRGLSTFSSLANARGLCTAHBGLVFOHDTSGOOLGRVLDVLRQ 262
 OY 263 LAQAGATVVVVFSSRQLARVFESVYVLTNLGKVVVASEAMALSRHITGPGIORIGVL 322
 DB 263 VNOSKYQVVVLFASARAVYSLEFSYSIHGSLSPKVVVASESWLTSDLVMTLPINARGYVL 322
 OY 323 GVAIQRAVPGILKAEEAYARADKEAPRCHGKSCSS-----NOLCRECOAF 370
 DB 323 GFLQGRALLPERSHYETHLAL--AADP---AFCASLNAELDEEHVWGRCRCDI 375
 OY 371 MAHTMPKLKAFSSSA-----YNAYRAVYVAHGLHQLGASGACS-GRGYVPQOLLE 423

```

Db 376 MLONLSSGLLONLSAGOLHHQIFATYAAYSVAGALHNHTLQCNVSHCHVSEHVLPMOLLE 435
QY 424 QIKRVHFLHKDYAFNDNDPLSSYNIITAMDNGPKMTFTVLGSSWSPVOLINNETKI 483
Db 436 NMYNMSFHARDLTLOFAEGNVMEYDLKMMWQSPFVLTHTVGTFTNGT---LQLOQSKM 492
QY 484 QMHGKDNQPKSVSCSDCLSGHQRVYTFHHCCEPCVPCAGTFLNKSDDLYRCOPCGKEE 543
Db 493 YWPG--NQVPVSCSRCKGQVRKVGKFSCTCDVCDAKASRKRPRDFTCTPCMODQ 550
QY 544 WABEGSOTCEPRRTVFLALREHNSWVLAANTLLLLLTAGLFAHMLDTPVYRSAG- 602
Db 551 WSPKSTACLPRRPKFLAMGEPVYVLLDLCVLTGLALALGSLVHMDSPVLQASGGS 610
QY 603 RLCE--LMLGSLAAGSSLYGFGEPRPACILRQALFGLTFLGCLVYRSQOLIIF 660
Db 611 QFEGGLICGLFCL--SVLLFPGRPSSASCLAQOPAMHRLPTECLSTFLQAAETV- 665
QY 661 KFSKVPTEFYHAWYQNHGAGLE---VMISSAQLLICTLWLVWTPLPAREYORFPHLY 716
Db 666 --ESELPLSNANMLCSYLRGLAMLVYLLATFVEAALCAWTLIAFPPEVYTDMSVLPREV 723
QY 717 MLECTETNSLGLFIATLYNGLSLISAFACSYLGKDPENYNEAKCVTSLLENFVSWIAF 776
Db 724 LEHCHVSWSLGLVHTTNAMLAFLCGLTFLVOSOPGRYRARGLTFAMLAFTWVSF 783
QY 777 FTTHSVYDGYLPANMAGLSSSGFGYFLPKCYVILCRPDLNTEHF 827
Db 784 VPLIANVOVAQOPAVOMGAILVCALGILVTHLPKCYVILMLPKLNTQEFF 834

```

RESULT 13

```

Q91VA4 PRELIMINARY; PRT; 858 AA.
ID 091VA4;
AC 091VA4;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PUTATIVE TASTE RECEPTOR (CANDIDATE TASTE RECEPTOR T1R3) (PUTATIVE
DE SWEET TASTE RECEPTOR TYPE 1 MEMBER 3).
GN T1R3 OR TAS1R3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A..
RC STRAIN=C57BL/6N; TISSUE-BRAIN;
RA Kitagawa M.;
RT "a putative taste receptor."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A..
RC STRAIN=C57BL/6J;
RX MEDLINE=21219400; Pubmed=11319557;
RA Montmayeur J.P., Liberles S.D., Matsunami H., Buck L.B.;
RT "A candidate taste receptor gene near a sweet taste locus."
RL Nat. Neurosci. 4:492-498(2001).
RN [3]
RP SEQUENCE FROM N.A..
RC STRAIN=C57BL/6J;
RA Bachmanov A.A., Li X., Reed D.R., Ohmen J.D., Li S., Chen Z.,
RA Torzof M.G., deJong P., Wu C., West D.B., Chatterjee A., Ross D.A.,
RA Beauchamp G.K.;
RT "Positional cloning of the mouse saccharin preference (Sac) locus."
RL Chem. Senses 0:0-0(2001).
RN [4]
RP SEQUENCE FROM N.A..
RC TISSUE=CIRCUMVALLATE PAPILLAE;
RX MEDLINE=21222875; Pubmed=11322794;
RA Kitagawa M., Kusakabe Y., Miura H., Niomiya Y., Hino A.;
RT "Molecular Genetic Identification of a Candidate Receptor Gene for

```

```

RT Sweet Taste."
RL Biochem. Biophys. Res. Commun. 283:236-242(2001).
DR EMBL; AB055708; BAB62952.1; -
DR EMBL; AF337039; AAK39436.1; -
DR EMBL; AF311386; AAL08425.1; -
DR EMBL; AB049994; BAB47181.1; -
KW Receptor.
SO SEQUENCE 858 AA; 94561 MW; 4E4168279DB478F8 CRC64;

```

```

Query Match 25.5%; Score 1147; DB 11; Length 858;
Best local similarity 32.6%; Pred. No. 6.8e-88;
Matches 271; Conservative 150; Mismatches 356; Indels 54; Gaps 16;

```

```

QY 29 SPDTFLPGDYLALFLPHS---GCLQVRHREPVLCDRSCSFNKHGHLQAMALGYEE 85
Db 26 SQOFRAGDYLLGLFPLGSTEATLNTORTNPSIPCNR---FSPGLFLMAMMAVEE 82
QY 86 INNSTALPLNTLLGQYLDVCSDS--ANYATLRYLSLPGOHIELOGDLHYSPVLAVI 144
Db 83 INNGSALLPGLRLGDLEDFTCSEPVYTMKSSLMFLAKVSGSIAAYCVYQYQPRVLAVI 142
QY 145 GPDSINRAATTAALISPLVPMISTAASETLSVKRQYPSFLRTIPNDKYOVETWYLLQ 204
Db 143 GPHSESLALITGKEFSEFLMPQVSYASMDRLSDRETPSEFRTVPSPDROVLQAVYTLQ 202
QY 205 KFGMTWISLVSDDYGGOLGOALENQTGGICIAFQDIMPFSNOVDE--RMQCLMRH 262
Db 203 NFSNMWVAALSDDDYGEGLSTPSSLANARGICIAHGGLVPQHTSQQLAKYLDVLRQ 262
QY 263 LAQAGATVVVVSRROLARVEESVYLTNLTKVWVASEANALSRHTGVGQIRIGVL 322
Db 263 VNQSKVQVYVLFASARAVYSLFYSYIHGGLSPKVVAVASEMLTSDLVMTLPILAVGVYL 322
QY 323 GVALQKRAVPGKAFEEYARADKEAPRPCHKSGKSS-----NQLCEQCAF 370
Db 323 GFLQRGALLPEFSSHVETHTAL--AADP---AFCAALNLELDEHYMGQRCPRCDI 375
QY 371 MAHMPKTKAFMSA-----YNAVYRAVVAAGLHOLLCASGACS--RGVYPMOLLE 423
Db 376 MLONLSSGLLONLSAGOLHHQIFATYAAYSVAGALHNHTLQCNVSHCHVSEHVLPMOLLE 435
QY 424 QIKRVHFLHKDYAFNDNDPLSSYNIITAMDNGPKMTFTVLGSSWSPVOLINNETKI 483
Db 436 NMYNMSFHARDLTLOFAEGNVMEYDLKMMWQSPFVLTHTVGTFTNGT---LQLOQSKM 492
QY 484 QMHGKDNQPKSVSCSDCLSGHQRVYTFHHCCEPCVPCAGTFLNKSDDLYRCOPCGKEE 543
Db 493 YWPG--NQVPVSCSRCKGQVRKVGKFSCTCDVCDAKASRKRPRDFTCTPCMODQ 550
QY 544 WABEGSOTCEPRRTVFLALREHNSWVLAANTLLLLLTAGLFAHMLDTPVYRSAG- 602
Db 551 WSPKSTACLPRRPKFLAMGEPVYVLLDLCVLTGLALALGSLVHMDSPVLQASGGS 610
QY 603 RLCE--LMLGSLAAGSSLYGFGEPRPACILRQALFGLTFLGCLVYRSQOLIIF 660
Db 611 QFEGGLICGLFCL--SVLLFPGRPSSASCLAQOPAMHRLPTECLSTFLQAAETV- 665
QY 661 KFSKVPTEFYHAWYQNHGAGLE---VMISSAQLLICTLWLVWTPLPAREYORFPHLY 716
Db 666 --ESELPLSNANMLCSYLRGLAMLVYLLATFVEAALCAWTLIAFPPEVYTDMSVLPREV 723
QY 717 MLECTETNSLGLFIATLYNGLSLISAFACSYLGKDPENYNEAKCVTSLLENFVSWIAF 776
Db 724 LEHCHVSWSLGLVHTTNAMLAFLCGLTFLVOSOPGRYRARGLTFAMLAFTWVSF 783
QY 777 FTTHSVYDGYLPANMAGLSSSGFGYFLPKCYVILCRPDLNTEHF 827
Db 784 VPLIANVOVAQOPAVOMGAILVCALGILVTHLPKCYVILMLPKLNTQEFF 834

```

RESULT 14

Q923K0

ID 0923K0 PRELIMINARY: PRT: 858 AA.
 AC 0923K0:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SWEET TASTE RECEPTOR T1R3.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBITaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SMR/J;
 RA Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
 RA Zuker C.S.;
 RT "Mammalian Sweet Taste Receptors."
 RL Cell 0:0-0(2001).
 DR EMBL, AY032621; AAK51602.1; -
 KW Receptor.
 SQ SEQUENCE 858 AA; 94530 MW; 686A7E524267796A CRC64;

Query Match 25.4%; Score 1141; DB 11; Length 858;
 Best Local Similarity 32.6%; Pred. No. 2.2e-87;
 Matches 271; Conservative 149; Mismatches 357; Indels 54; Gaps 16;

29 SPDFLLPGDYLLAGLFPPLHS---GCLQVHRPEVTLCDRSCSFNEHGYLFOAMRLGVEE 85
 26 SQGFNAQGDYILGLGFLPLGSTEALNORTOPNSILCNF---FSPGLFLAAMKAAVE 82
 86 INNSTALLPNTLGYOLYDVCSDS-ANYATLRLVSLPGQHIEGLDILHSPYLAIV 144
 83 INNGALLPGLTGLDYDLEPTCEPVTMKSSIMFLAKVSGSIAAVCYTOYOPRYLAIV 142
 145 GPDSTNRATTAALLSPFLYPMISAASETISVKROYSEFRTIPNKXYQETWVLLQ 204
 143 GHSEELALITGKFSFPLMPOVYSASMDRLSDRETFEFPYPSDVOLOAVVTLQ 202
 205 KFGWTWISLVSSDDYQOLGVALENOATGQICIAFKDIMPFSAOVDE--RMOCIMRH 262
 203 NFSMWVAALGSDDDYGRGLSIFSSLANARICIAHBEVLQPHDTSQGLKVLVDLQ 262
 263 LAQAGATVYVYSSKQLARVFFESVYLTNLGKVVAVASAMLSRHITGPEIGIOMVL 322
 263 VNOSVVOYVYVFAARAVYSIFSYSIHGLSPKVVAVASESWLTSOLVMTLPIMARGLVL 322
 323 GVAIGKRAVPGKAFEEYARADKEAPRCHGSCSS-----NOLCEGCAQF 370
 323 GFLRGALLPERSHYETHLAL---AADP---AFCASLNAELDLBEHYMGORPCDDI 375
 371 MAHTPKLKAFSSMSA-----YNAVRAVYVAHGLHOLLGASGASG--RGHYVPMOLLE 423
 376 MLONLSGLQLSLAGOLHGHOFATYAAVYVAQALHNTLOCNVSHCHSEHYVLPOLLE 435
 424 OIHKHFLHLKDYAFNDNRDPLSYNITLADWNGPKTFTYLGSSITSPVOLINETKI 483
 436 NMYNNSFARDLTLQFDEGNVDMEDLKMVWOSPRTVLTHTVGFNGT---LOLOOSAM 492
 484 OMHGKDNQVPSKSSDCEGHQVAVTGFPHCCFECVPGCACTPLNKSPLYRCOPGKEE 543
 493 YWPG--NOVPVSGCRCKDGOVRRYKGFHSCCTVCDCKASIKRHPDDFTCTPQNDQ 550
 544 WAPESQTCFPTVVELALREHTSVLLAANTLLDLLLGTAGLFAMHLDTPVVSAG- 602
 551 WSPKSTACLPRRKFLAMGERVYVLSLLLCVLVGLALALAGLSVHHNDSPVQASGS 610
 603 RLCF--LMLGSLAAGSGSYGFFGEPTRPACLLROALFALGTTPLSCITVRSFOLLITF 660
 611 QCFELICIGFCL---SVLLFPGRRSSASCLAOOPMAHLPLTGLCISTLFLAAGTFV-- 665
 661 KFTSTVPTFYHAMVONHAGLF---VMISSAOLLLCITLVVVTPLPARYQGFPHVY 716

DB 666 --ESELPLSMANWLCSYLGLNAMLVLSATPVEALCAMYLAPPEVYDTMSVLPTEV 723
 QY 717 MLECFETNSLGLFLAVLYNGLSISAFACSYLGKDPENYNEAKCYTFLTFNFWIMAF 776
 DB 724 LEHCYRSMVSLGVLHITAMALFLCFLGTFVQSPGKRYNARGLTFMAYLFTWVSF 783
 QY 777 FTTASVYDGKYLPAANMMAGLSLSSGFGGYFLPKCYVILLCRPDLNSTEHEF 827
 DB 784 VPLLANVQAVQPAVQMGAILVCAIGLITVFLPKCYVLLMPLKNTQGEFF 834

RESULT 15
 ID 073635 PRELIMINARY: PRT: 940 AA.
 AC 073635:
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CALCIUM2+ SENSING RECEPTOR.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 CC Tetraodontidae; Takifugu.
 OX NCBITaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96226788; PubMed=9560249;
 RA Naito T., Saito Y., Yamamoto J., Nozaki Y., Tomura K., Hazama M.,
 RA Nakanishi S., Brenner S.;
 RT "Putative pheromone receptors related to the Ca2+-sensing receptor in
 RT fugu."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:5178-5181(1998).
 DR EMBL: AB008857; BAA26122.1; -
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR InterPro: IPR000651; RasGEFN.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECPEP_F3_1; 1.
 DR PROSITE: PS0259; G_PROTEIN_RECPEP_F3_4; 1.
 KW Receptor.
 SQ SEQUENCE 940 AA; 105814 MW; 06DAB7803B6878B3 CRC64;

Query Match 23.6%; Score 1060.5; DB 13; Length 940;
 Best Local Similarity 30.8%; Pred. No. 1.5e-80;
 Matches 262; Conservative 155; Mismatches 361; Indels 73; Gaps 19;

34 LPQDYLLAGLFPPLHSGCL---QVHRPEVTLCDRSCSFNEHGYLFOAMRLGVEEINNS 89
 29 MTGDLILGLGLFPIHGISCKDENLARPSTKCVR---FNFGFRLQAMVPAIEIRINS 85
 90 TALLPNTITGOLYDVCSANVYATLRLVSLPGQHIELOG-----DLHSPYLAIV 144
 86 SSLPNTITGYRIFPTCNVSK--ALEATLSVAONKIDSLNDEFQNTDHPATIAV 143
 145 GPDSTNRATTAALLSPFLYPMISAASETISVKROYSEFRTIPNDYQETWVLLQ 204
 144 GAAGSVAIVAVANLSTFLTIPOISTAASSRLSNKNQYSPKRTIPTDHOCTAADYIE 203
 205 KFGWTWISLVSSDDYQOLGVALENOATGQICIAFKDIMPFSAOVDERMOCLMRHLA 264
 204 YQWMMVAVASDDDYGRGRIEKFEMERDICTHNLNLI--SQYFECETKALVDRIE 261
 265 QAGATVYVYSSKQLARVFFESVYLTNLGKVVAVASAMLSRHITGVPGIORIGVLCV 324
 262 NSTAVIVVFAAGPDIPLIKEMVRNITDRIMLASEMASSSLAKPEYLDVVGTTIF 321
 325 AIOKRAVPGKLAFF-----EEYARADKCAPR--PCHKSGSCSSNOL 363
 322 VLKAGNIFQFREFLQOVQPKRGSNHEFYAEFEETFNCTLEDSPRLQSENGS--DSFRL 380

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 10:18:37 ; Search time 25.2 Seconds

(without alignments)
1292.188 Million cell updates/sec

Title: US-09-819-946-2

Perfect score: 4493

Sequence: 1 MLCTARLVGLQLISCCWA.....NSTEHFQASIDPYTRRCGST 841

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1142	25.4	1078	1	CASR_HUMAN
2	1126.5	25.1	1085	1	P5180 homo sapien
3	1123	25.0	1079	1	CASR_BOVIN
4	1105.5	24.6	1079	1	CASR_MOUSE
5	741.5	16.5	1212	1	MGR5_HUMAN
6	727	16.2	1199	1	MGR1_RAT
7	727	16.2	1203	1	MGR5_RAT
8	719	16.0	1194	1	MGR1_HUMAN
9	708.5	15.8	912	1	MGR4_RAT
10	704.5	15.7	912	1	MGR4_HUMAN
11	692.5	15.4	872	1	MGR2_RAT
12	688	15.3	879	1	MGR3_RAT
13	684	15.2	877	1	MGR3_HUMAN
14	666.5	14.8	872	1	MGR2_HUMAN
15	666.5	14.8	877	1	MGR6_HUMAN
16	664	14.8	871	1	MGR6_RAT
17	654	14.6	908	1	MGR8_RAT
18	652	14.5	908	1	MGR8_MOUSE
19	650	14.5	908	1	MGR8_HUMAN
20	626.5	13.9	976	1	MGR_DROME
21	619.5	13.8	915	1	MGR7_HUMAN
22	619.5	13.8	915	1	MGR7_RAT
23	590.5	13.1	999	1	MGR1_MOUSE
24	186.5	4.3	960	1	MGR1_MOUSE
25	186.5	4.2	940	1	MGR2_RAT
26	185	4.1	961	1	MGR1_HUMAN
27	183.5	4.1	991	1	MGR1_RAT
28	172.5	3.8	941	1	MGR2_HUMAN
29	163	3.6	938	1	MGR2_HUMAN
30	158	3.5	938	1	MGR1_RAT
31	156	3.5	938	1	MGR1_MOUSE
32	144	3.2	1103	1	CYGF_BOVIN
33	135	3.0	1108	1	CYGF_HUMAN

34	124	2.8	986	1	CYGR_ARBP	P11528 arabacia pun
35	112.5	2.5	1125	1	CYGS_STRPU	P16065 strongyloce
36	110.5	2.5	908	1	GLK2_HUMAN	O13002 homo sapien
37	109	2.4	908	1	GLK2_RAT	P42260 rattus norv
38	108.5	2.4	540	1	ANPC_HUMAN	P17342 homo sapien
39	108.5	2.4	1108	1	CYGF_RAT	P51842 rattus norv
40	108	2.4	1350	1	XFIN_XENLA	P08045 xenopus lae
41	107.5	2.4	803	1	2226_HUMAN	O9ayt6 homo sapien
42	107	2.4	889	1	GLK2_MOUSE	P39087 mus musculu
43	106.5	2.4	545	1	NUSM_ALBGO	P48918 albinaria c
44	106.5	2.4	557	1	GHT4_SCHPO	O59932 schizosacch
45	106	2.4	487	1	BCH2_RHOGE	O9jpb9 rhodocyclus

ALIGNMENTS

RESULT 1
ID CASR_HUMAN STANDARD: PRT: 1078 AA.
AC P41180: Q13912; Q16379; Q16108; Q16109; Q16110:
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Extracellular calcium-sensing receptor precursor (Casr) (Parathyroid
DE Cell calcium-sensing receptor).
GN CASR OR GPRC2A OR PCARL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pearce S.H.S., Thakker R.V.;
RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE=Parathyroid;
RC MEDLINE=95279439; PubMed=7759551;
RA Garrett J.E., Capuano I.V., Hammerland L.G., Hung B.C., Brown E.M.,
RT Hebert S.C., Nemeth E.F., Fuller F.;
RL "Molecular cloning and functional expression of human parathyroid
RL calcium receptor cDNAs.";
RN J. Biol. Chem. 270:12919-12925(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE=Kidney;
RC MEDLINE=95408281; PubMed=7677761;
RA Aida K., Kohshi S., Tawata M., Onaya T.;
RT "Molecular cloning of a putative Ca(2+)-sensing receptor cDNA from
RT human kidney.";
RN Biochem. Biophys. Res. Commun. 214:524-529(1995).
RN [4]
RP SEQUENCE FROM N.A.
RA MEDLINE=96343808; PubMed=8756555;
RC Freichel M., Zink-Lorenz A., Hollrosch A., Hafner M., Flockert V.,
RA Rave F.;
RT "Expression of a calcium-sensing receptor in a human medullary
RT thyroid carcinoma cell line and its contribution to calcitonin
RT secretion.";
RN Endocrinology 137:3842-3848(1996).
RN [5]
RP SEQUENCE OF 643-908 FROM N.A.
RA MEDLINE=96193893; PubMed=8613532;
RC Biale D.D., Ratnam A., Mauro T., Harris J., Pillal S.;
RT "Changes in calcium responsiveness and handling during keratinocyte
RT differentiation. Potential role of the calcium receptor.";
RN J. Clin. Invest. 97:1085-1093(1996).
RN [6]
RP VARIANTS FHH GLN-165; LYS-297 AND TRP-795.
RA MEDLINE=94094324; PubMed=7916660;
RC Pollak M.R., Brown E.M., Chou Y.-H.W., Hebert S.C., Marx S.J.,
RA Steinmann B., Levi T., Seidman C.E., Seidman J.G.;

RT "Mutations in the human Ca(2+)-sensing receptor gene cause familial
RT hypocalcemic hypercalcaemia and neonatal severe
RT hyperparathyroidism.";
RL Cell 75:1297-1303(1993).
[7]
RP VARIANT ADH ALA-127.
RX MEDLINE=95179179; PubMed=7874174;
RA Pollak M.R., Brown E.M., Estep H.U., McLaine P.N., Kifor O., Park J.,
RA Hebert S.C., Seidman C.E., Seidman J.G.;
RT Autosomal dominant hypocalcaemia caused by a Ca(2+)-sensing receptor
RT gene mutation.";
RL Nat. Genet. 8:303-307(1994).
[8]
RP VARIANTS FHH MET-62; CYS-66; MET-138; GLU-143 AND GLN-227.
RX MEDLINE=95243222; PubMed=7726161;
RA Chou Y.-H.W., Pollak M.R., Brandt M.L., Toss G., Argyrist H.,
RA Atkinson A.B., Papapoulos S.E., Marx S., Brown E.M., Seidman J.G.,
RA Seidman C.E.;
RT "Mutations in the human Ca(2+)-sensing receptor gene that cause
RT familial hypocalcemic hypercalcaemia.";
RL Am. J. Hum. Genet. 56:1075-1079(1995).
[9]
RP SEQUENCE OF 1-61 FROM N.A., AND VARIANT FHH ALA-39.
RX MEDLINE=95403641; PubMed=7673400;
RT "Familial hypocalcemic hypercalcaemia associated with mutation in the
RT human Ca(2+)-sensing receptor gene.";
RL J. Clin. Endocrinol. Metab. 80:2594-2598(1995).
[10]
RP VARIANTS NSHPT LEU-227 AND TYR-598.
RX MEDLINE=96292293; PubMed=8675635;
RA Pearce S.H.S., Trump D., Wooding C., Besser G.M., Chew S.L.,
RA Grant D.B., Heath D.A., Hughes I.A., Patterson C.R., Whyte M.P.,
RA Thakker R.V.;
RT "Calcium-sensing receptor mutations in familial benign hypercalcaemia
RT and neonatal hyperparathyroidism.";
RL J. Clin. Invest. 96:2683-2692(1995).
[11]
RP VARIANTS ADHP THR-116; HIS-681 AND SER-806, AND VARIANT SER-851.
RX MEDLINE=96311554; PubMed=8733126;
RA Baron J., Winer K.K., Yanovsky J.A., Cunningham A.W., Laue L.,
RA Zimmerman D., Cutler G.B. Jr.;
RT "Mutations in the Ca(2+)-sensing receptor gene cause autosomal
RT dominant and sporadic hypoparathyroidism.";
RL Hum. Mol. Genet. 5:601-606(1996).
[12]
RP VARIANT FHH ARG-174.
RX MEDLINE=97442275; PubMed=9298824;
RA Ward B.K., Stuckey B.G.A., Gutteridge D.H., Laine N.G., Pullan P.T.,
RA Ratajczak T.;
RT "A novel mutation (L174R) in the Ca2+-sensing receptor gene
RT associated with familial hypocalcemic hypercalcaemia.";
RL Hum. Mutat. 10:233-235(1997).
[13]
RP VARIANT FHH GLU-557.
RX MEDLINE=21603857; PubMed=11762699;
RA Nakayama T., Minato M., Nakagawa M., Soma M., Tobe H., Aoi N.,
RA Kosuge K., Sato M., Ozawa Y., Kamatsue K., Kokubun S.;
RT "A novel mutation in Ca2+-sensing receptor gene in familial
RT hypocalcemic hypercalcaemia.";
RL Endocrine 15:277-282(2001).
-1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC
CC -1- TISSUE SPECIFICITY: FOUND IN KIDNEY, BUT NOT IN BRAIN, LUNG,
CC LIVER, HEART, SKELETAL MUSCLE, OR PLACENTA.
CC
CC -1- DISEASE: DEFECTS IN CASR ARE A CAUSE OF FAMILIAL HYPOCALCAEMIC
CC HYPERCALCAEMIA (FHH) AND NEONATAL SEVERE HYPERPARATHYROIDISM
CC (NSHPT), TWO INHERITED CONDITIONS CHARACTERIZED BY ALTERED CALCIUM

CC HOMEOSTASIS. THE MUTATIONS REDUCE THE ACTIVITY OF THE RECEPTOR.
CC FHH AFFECTED INDIVIDUALS EXHIBIT MILD OR MODERATE HYPERCALCAEMIA,
CC RELATIVE HYPOCALCAEMIA, AND INAPPROPRIATELY NORMAL PTH LEVELS. IN
CC CONTRAST, NSHPT IS A RARE AUTOSOMAL RECESSIVE LIFE-THREATENING
CC DISORDER CHARACTERIZED BY VERY HIGH SERUM CALCIUM CONCENTRATIONS,
CC SKELETAL DEMINERALIZATION, AND PARATHYROID HYPERPLASIA. IN SOME
CC INSTANCES NSHPT HAS BEEN DEMONSTRATED TO BE THE HOMOZYGOUS FORM OF
CC FHH.
CC
CC -1- DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT
CC HYPOCALCAEMIA (ADH) IN WHICH THE RECEPTOR IS ACTIVATED AT SUBNORMAL
CC CA(2+) LEVELS.
CC
CC -1- DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT
CC HYPOPARATHYROIDISM (ADHP). ADHP IS CHARACTERIZED BY HYPOCALCAEMIA
CC AND HYPERPHOSPHATEMIA DUE TO INADEQUATE SECRETION OF PARATHYROID
CC HORMONE. SYMPTOMS ARE SEIZURES, TETANY AND CRAMPS.
CC
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC
CC EMBL; X81086; CAA56990.1; -;
CC EMBL; U20759; AAA86503.1; -;
CC EMBL; U20760; AAA86504.1; -;
CC EMBL; D50855; BAA09453.1; -;
CC EMBL; S83176; AAB46873.1; -;
CC EMBL; S79217; AAB35262.2; -;
CC EMBL; S68032; AAB29413.2; ALT. SEQ.
CC EMBL; S68033; AAB29414.1; -;
CC EMBL; S68036; AAB29415.1; -;
CC EMBL; S81755; AAD14370.1; -;
CC GCRDB; GCR_1337; -;
CC GCRDB; GCR_1874; -;
CC GCRDB; GCR_2012; -;
CC GCRDB; GCR_2013; -;
CC GCRDB; GCR_2696; -;
CC GCRDB; GCR_2697; -;
CC MIM; 601199; -;
CC MIM; 145980; -;
CC MIM; 239200; -;
CC MIM; 601198; -;
CC InterPro: IPR001828; ANF_receptor.
CC InterPro: IPR000337; GPCR_Mgr.
CC Pfam: PF00003; 7tm_3; 1.
CC Pfam: PF01094; ANF_receptor; 1.
CC PRINTS; PR00248; GPCR_MGR.
CC PROSITE; PS00979; G_PROTEIN_REC_P3_1; 1.
CC PROSITE; PS00980; G_PROTEIN_REC_P3_2; 1.
CC PROSITE; PS00981; G_PROTEIN_REC_P3_3; 1.
CC PROSITE; PS00259; G_PROTEIN_REC_P3_4; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Disease mutation; Alternative splicing; Polymorphism.
CC
CC SIGNAL 1 19
CC CHAIN 20 1078
CC DOMAIN 20 612
CC TRANSMEM 613 635
CC TRANSMEM 636 649
CC TRANSMEM 650 670
CC TRANSMEM 671 681
CC TRANSMEM 682 700
CC TRANSMEM 701 724
CC TRANSMEM 725 745
CC TRANSMEM 746 769
CC TRANSMEM 770 792
CC TRANSMEM 793 805
CC TRANSMEM 806 828
CC TRANSMEM 829 836
CC TRANSMEM 837 862
CC DOMAIN 863 1078
CC
CC I (POTENTIAL).
CC CYTOPLASMIC (POTENTIAL).
CC II (POTENTIAL).
CC EXTRACELLULAR (POTENTIAL).
CC III (POTENTIAL).
CC CYTOPLASMIC (POTENTIAL).
CC IV (POTENTIAL).
CC EXTRACELLULAR (POTENTIAL).
CC V (POTENTIAL).
CC CYTOPLASMIC (POTENTIAL).
CC VI (POTENTIAL).
CC EXTRACELLULAR (POTENTIAL).
CC VII (POTENTIAL).
CC CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1079 AA; 120867 MW; D7664550361F9736 CRC64;

Query Match 25.0%; Score 1123; DB 1; Length 1079;
 Best Local Similarity 30.9%; Pred. No. 4,1e-74;
 Matches 275; Conservative 162; Mismatches 360; Indels 92; Gaps 22;

16 SCC---MAFACHSTRESSPD--FTLPGLDLAGLPLHSGCL---QVHREPVTLCDSC 66
 5 SCCALALLAHSSAYGPDRORAKKGDILLGLGFPHGVAKKODDLSPRESEVICR-- 62
 67 SFNEGHYHLPQAMRLGVEINNSTALLPNTLGYQLYDVCSDSANVYATLRLSLPGQH 126
 63 -YNRGFRMLQAMTFAIEINSSPLPNTLGYRIPTCTMYSK--ALEATLSFVAQNK 119
 127 IELDG-----DLHSPVYAVIGPDSINRAATTAALLSPFLVPMISTAASSETLSVRQ 181
 120 IDSLNIDFCNCSEHPISTIAVVGATGSGVSTAVANLGLFYIPQVSYASSRLSNKQ 179
 182 YPSFLRTIPNDKYQVETWLLQKFGMTWISLVGSSDDYDGOGLVALENOATGQICIAF 241
 180 YKSEFLRTIPNDHQATMAADIEFRMWVGTIAADDDYGRGICKEFEAEFERICIDF 239
 242 KDIMPFAOVDERMOCLMRHLAAGATVYVVFSSROLARVFESVLTNLTKYVASE 301
 240 SELI--SQYSEDEEIQVVEYQNSTAKYIVVFSSGPDLEPLIKETVIRNITGRIMASE 297
 302 AMALSRHTVPGIORIGMVGVALQKRAVPGIKAF-----EAAVRAKDEARRP 351
 298 AMASSSLIAMPDEYFHVGGTIGFGIKAGQIPGFREFLOKVPKRSVHNGFAEFMEETFN 357
 352 CH-----KG-----SWCSSNOLRECOAFMAHMKLAKARMSA 386
 358 CHLGEGAGPLPYDTFVNSHEGGRRLNLSSTAFRLCTGDEINISVETPYMDYHLLAIS 417
 367 YNAVRAVYVAHGLHQLIGC-----ASGACSR-GRAYPMOLLEQIHKVHFLHK-DTV 437
 418 YNVYLAAYVIAHALODITYCLPGRCLFTNGSCADIKKVEAOVLKHLRLHNTNNMGROY 477
 438 AFNPNRDLSSYNTIATMMNNGPKTFTYLGSTSPV-----QLNNEKTIOMH 486
 478 TFDECGDLVGNSTIINMHLSPED-----GSTVEKEVGYVNVYAKKGERLFTNEKILMS 531
 487 GKDNQVPSKVSQDCLLEGHOR-VVYGFHHCFCFECVPCGAGTFLNKSRLDYRCOPCKEEMA 545
 532 GFSRVPFNSNRDQOATRGIIIEGPTCECEPCDEGSEGTDSACDKCDDRWIS 591
 546 PEGSOTCEPRVYVLALEHRTSWVLLAANTLLLLICTAGLEFANHLDTPVVRSAGRLC 605
 592 NENHTSIAKEIEFLAWEPEGIALTLFAVLGIFLTAFLVGIFFRMTPIYKATNRELS 651
 606 FLMGSLAAGSGSLGEGEPTRPACILROALFALGFTILFISCLTVYRSQOLIIRKFTK 665
 652 YLLFLSLCCFSSSLFFEGEPDWTCTRLQPAFGISFVLCISCLIVKTNRLVLEPE-AK 709
 666 VPTFVH-AMVONHGAGLFVMTSSAQLLICTLWLVVMTPLPAREYQRPPLHMLCETETN 724
 710 IPTSPHRKMWGLINQFLVLCFPMQILICIMLTATAPSSRNHLEDEIIFICHGGS 769
 725 --SLGFIATLYNGLISLISAFACSTYLGKDLDPENYNEACVTFSLLENFVSWIAFTT-AS 781
 770 LMAISLIG--YTCLLAICEFFFAKSRKLEPENNEAKIFTESMLIFFIWMISFIPAYAS 827
 782 VYDGYVLPAAAMMAGLSSLSGFGYFLPKCYVILCRPDLNSTEHPQAS 830
 828 TY-GKFSVAVEVIALIASFGILACIFPNKYIILFKPSRMTIEVRSS 875

RESULT 4
 CASR_MOUSE STANDARD; PRT: 1079 AA.
 AC Q9QY96; Q08968; Q88519; Q9QY95; Q9QZ08; Q9R1D6; Q9R1V2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (CasR) (Parathyroid
 DE Cell calcium-sensing receptor).
 GN CASR OR GPRC2A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090.
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RX MEDLINE=20092890; PubMed=10625652;
 RA Oda Y., Tu C.-L., Chang W., Crumrine D., Koemueves L., Mauro T.,
 RA Elias P.M., Bikle D.D.;
 RT "The calcium sensing receptor and its alternatively spliced form in
 RT murine epidermal differentiation.";
 RL J. Biol. Chem. 275:1183-1190(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC STRAIN=BLACK SWISS X 129/SV; TISSUE=Kidney;
 RX MEDLINE=20119279; PubMed=10652312;
 RA Pi M., Garner S.C., Flannery P., Spurney R.F., Quarles L.D.;
 RT "Sensing of extracellular cations in CasR-deficient osteoblasts.
 RT Evidence for a novel cation-sensing mechanism.";
 RL J. Biol. Chem. 275:3256-3263(2000).
 RN [3]
 RP SEQUENCE OF 256-600 FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC TISSUE=Epiphyseal cartilage;
 RX MEDLINE=20043955; PubMed=10579354;
 RA Chang W., Tu C., Chen T.-H., Komuves L., Oda Y., Pratt S.A.,
 RA Miller S., Shoback D.;
 RT "Expression and signal transduction of calcium-sensing receptors in
 RT cartilage and bone".
 RL Endocrinology 140:5883-5893(1999).
 RN [4]
 RP SEQUENCE OF 507-582 FROM N.A. (ISOFORM A).
 RC STRAIN=NMRI; TISSUE=Brain;
 RA Hildbrand J., Ammon H.P.T., Wahl M.A.;
 RT Submitted (May-1998) to the EMBL/Genbank/DBJ databases.
 RL [5]
 RP SEQUENCE OF 562-814 FROM N.A.
 RC TISSUE=Kidney;
 RA Moawad T.I., Ricciardi D.;
 RT Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 646-799 FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=97231187; PubMed=9076582;
 RA Quarles L.D., Hartle J.E. II, Siddhanti S.R., Guo R., Hinson T.K.;
 RT "A distinct cation-sensing mechanism in MC3T3-E1 osteoblasts
 RT functionally related to the calcium receptor".
 RL J. Bone Miner. Res. 12:393-402(1997).
 CC - FUNCTION: SENSES CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CC G-PROTEIN IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC - SUBCELLULAR LOCATION: Integral membrane protein.
 CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC - TISSUE SPECIFICITY: EPIDERMIS, KIDNEY AND CARTILAGE.
 CC - SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC or send an email to license@isb-sib.ch).

DR EMBL: AF110178; AAD28371.1; -

DR EMBL: AF110179; AAD28372.1; -

DR EMBL: AF128842; AAD40638.1; -

DR EMBL: AF068900; AAC19388.1; -

DR EMBL: AB027140; BAA77688.1; -

DR EMBL: AF002015; AAC53252.1; -

DR EMBL: AF159565; AAF00193.1; -

DR MGD: MGI:1351351; Casr

DR InterPro: IPR001828; ANF_receptor.

DR InterPro: IPR000337; GPCR_Mgt.

DR Pfam: PF00003; 7tm_3. 1.

DR Pfam: PF01094; ANF_receptor. 1.

DR PRINTS: PR00248; GPCRMR.

DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.

DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.

DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.

DR PROSITE: PS50259; G_PROTEIN_RECP_F3_4; 1.

DR G-protein coupled receptor; Transmembrane; Glycoprotein; signal;

KW Alternative splicing.

FT SIGNAL 1 19

FT CHAIN 20 1079

FT DOMAIN 20 612

FT TRANSMEM 613 635

FT DOMAIN 636 649

FT TRANSMEM 650 670

FT DOMAIN 671 681

FT TRANSMEM 682 700

FT DOMAIN 701 724

FT TRANSMEM 725 745

FT DOMAIN 746 769

FT TRANSMEM 770 792

FT DOMAIN 793 805

FT TRANSMEM 806 828

FT DOMAIN 829 836

FT TRANSMEM 837 862

FT DOMAIN 863 1079

FT CARBOHYD 90 130

FT CARBOHYD 130 261

FT CARBOHYD 261 287

FT CARBOHYD 287 386

FT CARBOHYD 386 446

FT CARBOHYD 446 468

FT CARBOHYD 468 488

FT CARBOHYD 488 541

FT CARBOHYD 541 594

FT CARBOHYD 594 537

FT VARSPLIC 451 537

FT VARSPLIC 45 45

FT CONFLICT 304 304

FT CONFLICT 410 410

FT CONFLICT 566 566

FT CONFLICT 595 595

FT CONFLICT 610 610

FT CONFLICT 814 814

FT CONFLICT 889 889

FT CONFLICT 906 909

FT CONFLICT 1057 1057

FT CONFLICT 1064 1064

FT CONFLICT 1076 1076

FT SEQUENCE 1079 AA; 120839 MW; AAF8BD8472736DE CRC64;

Query Match 24.6%; Score 1105.5; DB 1; Length 1079;

Best Local Similarity 30.5%; Pred. No. 7.7e-73;

Matches 271; Conservative 165; Mismatches 359; Indels 93; Gaps 22;

QY 14 LISCWAFACHSTRESSPD--FTLPGLDYLALFLPLHSGCL---QVHRHPEVTLCDRSCS 67

DB 10 LIALTM-----HSSAYGPDQRAQKKDIIILGGLPIHFHFGVAAKQDLKSRPESVCEIR--- 62

QY 68 FNEHGYHLEQAMRLGVEEINNSTALLPNTITGLGYQYDVCSDSANVATLRLVSLPGQHHI 127

DB 63 YNRCFRMLQAMFAIEEINSSPALLPNTITGLGRIPDTCNVSK--ALBATLSFYAQNKI 120

QY 128 ELQG-----DLHSTYVLAVIGPDSINAAATTAALLSFVPMISYAASETLSVKRQY 182

DB 121 DSIUNIDFNCSEHPIPTSTAVVANGSGSVTAVANLLGLFYIPQVYASSSRLLSKNPF 180

QY 183 PSFLRTIPNDKYQVETMVLVLLQFGWTYSLVGGSDPYQOLQVQALENOAGOCIGIAEK 242

DB 181 KSLRTIPNDENQATMADIIIEFRNNWGTIAADDYGRPIGEKREAEERDICIIDS 240

QY 243 DIMPSAQYGDERRQCLMHLAQAQAVVVFSSROLAAYEFESVYLNLTKVWVASEA 302

DB 241 EIL--SQYSDDEEIQVVEVIONSTAKIYVFSGDEPLLEIKETIRNITGRIMLASEA 298

QY 303 WALSRITGVPRGIRGMVGLVGIQKRAVPGIKAF-----EENARADKEAPRC 352

DB 299 WASSSLIAMEYFHVVGITIGFLKAGQIPGREFLQKVPKRSYHNGFAKEFEETENC 358

QY 353 H-----KG-----SMCSSNOLCRECOAFMAHTMPLKAFSMSAY 387

DB 359 HLDQAKGGLPVDTPYRSHBEGGNRLNSTAFRLPCTGDEMINSVETPMGEHLRISTY 418

QY 388 NAYRAYAVAHGLHQLLG-----ASGACSR--GRVYPMQLLEQIHKVPILHK--DTVA 438

DB 419 NYLAVYSIAHALQDIYTCLPGRGLETPNCSADIKKVEAMQVILKHLRLNFTNNMEQYT 478

QY 439 FNDNRPLSSYNIIMDNMGPMTFTVLGSSWSPV-----QINIEKTIQWNG 487

DB 479 FDECGDLVGNYSIINHLSPED-----GSIVKEVGYNNYAKKGERLEINIGKILMSG 532

QY 488 KDNQPKSVSCDCLLEGHOR--VVTGFHCCFECVPCAGTFPLKSLDYRCQPCGKEENAP 546

DB 533 FSRVEVFSKSCDQCGATKGIIEGEPICCFECVEPDEYSGEDDASCDKCPDPMFN 592

QY 547 EESQTCPEPTVPLALREHTSVLLAANTLLLLLGTAGLFAMHLDPVVSAGRGLCF 606

DB 593 EYVTSICIAEIEIFLMTPEFGIALTLFVAVGIFLFAVIGVIFKFRMPVIVATNRELSY 652

QY 607 IMGSLAAGSSLYGFGFEPTRPACLLROALFALGTITPLSCLYTRSPDLIIIFKSTV 666

DB 653 LILFSLCCFSSSLFTIGEPQWTCRLROPARGISFVLCISLKYKNVLYVE--AKI 710

QY 667 PTFYH--AMVQNHGAGLFVMISSAOLICLTWLVVWTPPLPAREYORFPLVMECTEN- 724

DB 711 PTFHKKMGMLQFLVFLCFEMQIVICIMLYTAPSPSYNNHEDEIIFITCHEGSL 770

QY 725 -SLGFLAFLYNGLSISAFACSYLQKDLPEYNEAKCYTESLRFVYSWIAFFT-ASY 782

DB 771 MALGSLIG--YTCLLAALICEFFAFKSRKLPENFENKAKFTFEMLIFFIWIISFIPAYAST 828

QY 783 YDQKYLPAANMAAGLSLSSGFGYPLPKCYILCPDPNSTEHOAS 830

DB 829 Y-GKVSAYEVAIILAASFGLLACIFPKVYIILKFRPSNLTIEVRSS 875

RESULT 5

ID MGR5_HUMAN STANDARD; PRI: 1212 AA.

AC PA1594;

DT 01-NOV-1995 (rel. 32, Created)

DT 01-OCT-1996 (rel. 34, Last sequence update)

DT 16-OCT-2001 (rel. 40, Last annotation update)

DE Metabotropic glutamate receptor 5 precursor.

GN GRM5 OR GPRC1E OR MGLUR5.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OC NCBI_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=94197696; PubMed=7908515;
 RA Matsumi R., Katsuki F., Yamamoto T., Nakamura K., Sugiyama H.:
 "Molecular cloning and the functional expression of two isoforms of
 human metabotropic glutamate receptor subtype 5.";
 RL Biochem. Biophys. Res. Commun. 199;1136-1143(1994).
 RN [2]
 RP REVISIONS.
 RA Katsuki F.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 860-952 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93343913; PubMed=7688218;
 RA Matsumi R., Katsuki F., Sugiyama H.:
 "A variant of metabotropic glutamate receptor subtype 5: an
 evolutionally conserved insertion with no termination codon.";
 RL Biochem. Biophys. Res. Commun. 194;622-627(1993).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
 CALCIUM SECOND MESSENGER SYSTEM AND GENERATES A CALCIUM-ACTIVATED
 CHLORIDE CURRENT.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 5A (SHOWN HERE) AND 5B; ARE
 PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER BY AN INSERTION OF
 32 RESIDUES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STROMEST, TO MGLU1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D28538; BAA05891.1; -;
 DR EMBL: D28539; BAA05892.1; -;
 DR EMBL: S64316; AAD13954.1; -;
 DR GCRDB: GCR_0761; -;
 DR GCRDB: GCR_1002; -;
 DR GCRDB: GCR_1003; -;
 DR GCRDB: GCR_1317; -;
 DR MIM: 604102; -;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgt.
 DR Pfam: Pf00003; 7tm_3; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECPT_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECPT_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECPT_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 1212
 FT DOMAIN 22 579
 FT TRANSSEM 580 602
 FT DOMAIN 603 616
 FT TRANSSEM 617 637
 FT DOMAIN 638 648
 FT TRANSSEM 649 667
 FT DOMAIN 668 693
 FT TRANSSEM 694 714
 FT DOMAIN 715 737
 FT TRANSSEM 738 759
 FT DOMAIN 760 772
 FT TRANSSEM 773 795
 FT DOMAIN 796 801
 FT TRANSSEM 802 827
 FT DOMAIN 828 1212
 FT CARBOHYD 88
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 METABOTROPIC GLUTAMATE RECEPTOR 5.
 I (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 II (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 III (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 IV (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 V (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 VI (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 VII (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 734 734 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 877 908 MISSING (IN ISOFORM 5A).
 SQ SEQUENCE 1212 AA; 132468 MW; A3C73606681C6A25 CRC64;
 Query Match 16.5%; Score 741.5; DB 1; Length 1212;
 Best Local Similarity 27.5%; Pred. No. 3.2e-46;
 Matches 238; Conservative 149; Mismatches 355; Indels 123; Gaps 33;
 QY 34 LPGLDLAGLPLHSGCLOVRRHREPV-FLDCSCSF--NEHQYLFQARLGVETINNST 90
 DB 31 MGDIIITIALSVH-----HQPTDYKHEKCAVGEQIQGVEMLHLEINSDP 83
 QY 91 ALLPNTILGOLYDVCSDSANYATRLVLSLPGHRIELQGLDLYS----- 137
 DB 84 TLLPNTILGCEIRDSCHWSA-----VALQSIETIRDSLSSEEGLVACVDGS 133
 QY 138 -----PTVLAVIGPDSTNRATTAALLSPFLVMISYASSSEFLSYKQYPSFLRTP 190
 DB 134 SSSFRSKRPVIGVIGPGSSVAIQVONLQENIQIAYSATSMDSKTLKRYMRVVP 193
 QY 191 NDKYGVETMVLIDKFGMTWISLVSDDYGGOLQVQALENQTGGGICIA--FKDIMFSS 248
 DB 194 SDAQARAMVDIVKRYNTYSAVITEGNYGSGHEAKRDKMSAKGICIAHSKITYSNAG 253
 QY 249 AOVGERMQLMRHLAQAGATVVVFSRQLARFESEVYLTLNGK--VWASEAMALSR 307
 DB 254 EQSPFKLLKTLKLSHPK--ARVACFCGEMVYRGLLMARRGLLAGEFLILGSDMADRY 311
 QY 308 HITGIPGIGRIGMVLGVAIQKRAVPGILAFEEAAYARADKEARPHKSKMCSNQLCDEC 367
 DB 312 DVT--DGYQR-EAVGIGITIKLOS-PDVKWFDDYL--KLRETNHRNRMF-----F 355
 QY 368 QAFNAHTMP-KLKAFSM--SSAYN-----AYRVAAYAHGLH--Q 402
 DB 356 QEFQHRFOCRLEGFPQENSKYKNTCSNLTIKTHVODSKKGFVIMATYSATYGLHMQ 415
 QY 403 LLGCA--SGACSR-----GRVYPMQLQIHKHFL-LHKDTVAPNDNRPLSSNYIIAM 454
 DB 416 MSICPGVAGLDAMKPIDGR-----KLESIMKTNFTGVSQDILFDENDQSGRVEIMNF 471
 QY 455 DMNGKMTFTVLGSGTMSFPVOLINETKIOMHGKNOYPKSYCSSDCLEGHORVY--TGFH 513
 DB 472 KEMGRDY-FDYINVGSMWNGELKMDDEY--WSKSN-IIRSVCSPECEKQIKYIRKEGV 528
 QY 514 HOCFECVPCAGTFLKNSDILVRCQPCGKEEMARBEQSQCPTVYVFLALRETSVLLAA 573
 DB 529 SCCWCTCPCKENEYV--FDEYTCRACQLGSWTDLTGDDLIPVYLLRMGDEPLAAVVF 586
 QY 574 NTLILLLLTAGLFAWMLDTPVVSAGGRCLFLMGLSIAAGSGSLYFGFEPTRPACIL 633
 DB 587 ACIGLATLFTVYVHTIIRDPVYVSSSRELCYIILLAGICGLYCLFCLIAKPKQIYCYL 646
 QY 634 ROALFALFTYPLSCLTVRSFOLIIF-----KSTKYPTFYHAYVONHAGLEFMISSA 688
 DB 647 ORIGIGLSPMSYSALVTKTRIAIRILAGSKKIKCTKPRFSACQALVIAFILLICI--- 703
 QY 689 AOLILCLMWVWMPPLAREVQRPFL--VMLEGETSISGLILAFNLNGLSIAFAPS 746
 DB 704 -QLGIIVALFMEPDMHDY--PSIREVYICMTTN-LGVVPLGNGLLISCTEY 758
 QY 747 YLGLDLPENYEAQVTSLLFENYSWIAFTTASVYDGKYLPAANMAAGLS-SLSGFG 805
 DB 759 EKTIRVAPNPNKAKIAFTMTTCLIMLAF---VPIYGSNKKITMCFVSLSATVALG 815
 QY 806 GYFLPKCYVILLCRPDLNSTEHPQAS 830
 DB 816 CMFVPKYVYIIIAKPRNVSATFTS 840

```

RESULT 6
ID MGR1_RAT STANDARD: PRT: 1199 AA.
AC P23385:
DT 01-NOV-1991 (rel. 20, Created)
DT 01-NOV-1991 (rel. 20, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Metabotropic glutamate receptor 1 precursor.
GN GRM1 OR GPRC1A OR MGLUR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI:taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=91156047; Pubmed=1847995;
RA Masu M., Tanabe Y., Tsuchida K., Shigemoto R., Nakanishi S.;
RT "Sequence and expression of a metabotropic glutamate receptor.";
RL Nature 349:760-765(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9202526; Pubmed=1656524;
RA Houamed K.M., Kujper U.L., Gilbert T.L., Haldeman B.A., O'Hara P.U.,
RA Mulvihill E.R., Almers W., Hagen F.S.;
RT "Cloning, expression, and gene structure of a G protein-coupled
RT glutamate receptor from rat brain.";
RL Science 252:1318-1321(1991).
RN [3]
RP ALTERNATIVE SPLICING (ISOFORM 1B).
RC TISSUE=Brain;
RX MEDLINE=92110002; Pubmed=1309649;
RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
RT "A family of metabotropic glutamate receptors.";
RL Neuron 8:169-179(1992).
RN [4]
RP ALTERNATIVE SPLICING (ISOFORM 1C).
RC TISSUE=Brain;
RX MEDLINE=93066232; Pubmed=1438218;
RA Pin J.-P., Maeder C., Prezeau L., Bockaert J., Heinemann S.F.;
RT "Alternative splicing generates metabotropic glutamate receptors
RT inducing different patterns of calcium release in Xenopus oocytes.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10331-10335(1992).
RN [5]
RP FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
RT MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
RT CALCIUM SECOND MESSENGER SYSTEM. MAY PARTICIPATE IN THE CENTRAL
RT ACTION OF GLUTAMATE AND LONG-TERM DEPRESSION IN THE CEREBELLUM.
RN [6]
RP THE HIPPOCAMPUS AND LONG-TERM DEPRESSION IN THE CEREBELLUM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A (SHOWN HERE), 1B AND 1C;
CC ARE PRODUCED BY ALTERNATIVE SPLICING. 1B AND 1C ARE C-TERMINALLY
CC TRUNCATED FORMS OF 1A.
CC -1- TISSUE SPECIFICITY: IS PREDOMINANTLY EXPRESSED IN CEREBELLAR
CC PURKINE CELLS, CA2-CA3 PYRAMIDAL CELLS OF THE HIPPOCAMPUS, AND
CC MITRAL AND TUFTED CELLS OF THE OLFACTORY BULB.
CC -1- MISCELLANEOUS: ACTIVATED BY QUISOVALATE > GLUTAMATE > IBOVENATE >
CC TRANS-1-AMINOCYCLOPENTYL-1,3-DICARBOXYLATE; INHIBITED BY
CC 2-AMINO-3-PHOSPHONOPROPIONATE.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLURS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X57569; CAA40799.1; -
DR EMBL: M61099; AAA19497.1; -

```

```

DR EMBL: S48085; AAB24138.1; -
DR PIR: S15362; S15362.
DR PIR: AA1939; AA1939.
DR GCRDB: GCR_0216; -.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF01094; ANF_receptor; 1.
DR PRINTS: PR00248; GPCR_MGR.
DR PROSITE: PS00979; G_PROTEIN_REC_P3_1; 1.
DR PROSITE: PS00980; G_PROTEIN_REC_P3_2; 1.
DR PROSITE: PS00981; G_PROTEIN_REC_P3_3; 1.
DR PROSITE: PS00982; G_PROTEIN_REC_P3_4; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family; Alternative splicing.
FT CHAIN 1..18
FT DOMAIN 19..592
FT TRANSMEM 593..615
FT DOMAIN 616..629
FT TRANSMEM 630..650
FT DOMAIN 651..661
FT TRANSMEM 662..680
FT DOMAIN 681..706
FT TRANSMEM 707..727
FT DOMAIN 728..750
FT TRANSMEM 751..772
FT DOMAIN 773..785
FT TRANSMEM 786..808
FT DOMAIN 809..814
FT TRANSMEM 815..840
FT DOMAIN 841..1199
FT DOMAIN 1014..1034
FT DOMAIN 1074..1080
FT DOMAIN 1126..1135
FT DOMAIN 1140..1199
FT CARBOHYD 98
FT CARBOHYD 223
FT CARBOHYD 397
FT CARBOHYD 515
FT VARSPLIC 887
FT VARSPLIC 906
FT VARSPLIC 907
FT VARSPLIC 888
FT VARSPLIC 898
FT SEQUENCE 1199 AA; 133235 MW; EEB5A04C050694B9F CRC64;

Query Match 16.2%; Score 727; DB 1; Length 1199;
Best Local Similarity 26.7%; Pred. No. 3; 6e-45;
Matches 230; Conservative 155; Mismatches 362; Indels 116; Gaps 32;

OY 34 LRGDYLAFLPLHSGCLVRRHREY-TECDSSCS--FNEHYHFLQARLTGVEEINNST 90
DB 41 MGDVYIAGLFVSH-----HQPRAEKYPERKCEGIRQYQIQVEAMFHLIDKINADP 93
OY 91 ALPNITLQYLYVDCSDSANVYATFLVSLPCQHHELGDL----- 134
DB 94 VLLPNTLSEIRDSCHWS-----SVALEGSEIFRISLISIRDEKDLNRCPLD 143
OY 135 -----HSPVYLVAVIGDSTNRKATTAALLSPVLPVISTASSETLSVKRQPSFLR 187
DB 144 GQTLPPGRTRKPIAGIYIGSSSSVAIQVONLQLETDIPDIAYSATSIDSLTKYFLR 203
OY 188 TTPNDKYQVEYVWLLQKSGMTWISLVGSSDDYGGOLGVOALNQMGQICIAFKDIMEP 247
DB 204 VVPSDILQRAMLDYKRIKNTYVSAVHTEGVSGMDAFELAAQBSCLTAHSD--KI 261
OY 248 SAQVDERAKQCLMRHLAQ--AGATVYVVFSSQRLARVFESYVLTNLTKGY--WVA 304
DB 262 YSNAGEKSFDRLLRKRLRELPKARVYVCCCEGMTYRGLLSABRRLGYVGESFLISD 321
OY 305 LSRHITGVPGIORIGKAVNLGVAIQKRAVPGLKAFEEAY--ARADKEAPRCPCHGSC 362

```

```

Db 322 DRDEIVEGEYEWANG---GITIKLOS--PEVRSFDDYFLKLDLNTNRNPFPEWQHRFQ 377
Qy 363 LCRCGQAFMAATM--PKLAF-----SMSSAY-----NAYRAYVAAGHL---HQLL 404
Db 378 -CR-----LPGHLEENPNEKVCCTGENSEENYVODSKMGVYINAIYAAHGLONNHVLC 432
Qy 405 GCASGACSR-----GRVYPMOLEQIHKVHFL-LHKDTVAFNDNRDPLSSNIATMDV-N 457
Db 433 PGHVGICDAMKPIDGR-----KLDFLIKSSFVSGEWEDEKDGADGRDINMLQYTE 488
Qy 458 GPKRTFVYLGSSVSPVOLINETKIQHGMKDNQVPSKSSCDLEGHQRYV-TGFHCC 516
Db 489 ANRDYVHVG--TWHEGLINIDYKIQMN-KSGMW-RSVCSEPCIKGQIKVIRKGEVSC 544
Qy 517 FECPGCGAGTFLNKSIDLXROPCKEEMAPRGSQTCPRRTVFLALRHTSWVLLAATL 576
Db 545 WICPACKEHEFV--ODETCRACDLGWWPMNMLTGCCEPIPRVLEWSDIESIIIAFEC 602
Qy 577 LLLLLTAGLFAHMLDTPVVSAGRLCFMLGLSLAGSGSLXGFCGEPRPACILROA 636
Db 603 GILVTLFVTLFVLYKRPDPVVKSSRELCTIILAGIFLGVCPTLLAKPTTSCYLORL 662
Qy 637 LEALGFTIFLSCLTVRSEQLIIF-----KESTVPTFYHAMVQNHGAGLFVMTSS--A 688
Db 663 LVGSAMCYSAVLTAKTRIRIARILAGSKKIKCTKRPRMSMAQ-----VIASILIS 715
Qy 689 AQLICLTLVYVTPPLAREQRPPLVMECTETNSIGFLIAPLNLGLSLISAFACYL 748
Db 716 VOLTLVTLIMEPPMLLSPISKEVYL--CMTSNLGVAAPVGYLMLTMSCTYVFK 773
Qy 749 GKDPENYNEAKCYTFSLEFVSMIAFPTASYDG-KYLPANMAGLSLSSGEGY 807
Db 774 TRNPANPNEKXYIAFTWYTCITLWAF---VPYFGSNKYIITTCRAVSLVVALGCM 830
Qy 808 FLPCYVILCRPDINSTEHQAS 830
Db 831 FTPKWTIIIAKPERNVSAPFTTS 853

RESULT 7
MGRS_RAT STANDARD; PRT; 1203 AA.
ID MGRS_RAT
AC P31424;
DT 01-JUL-1993 (Rel. 26, Last Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metabotropic glutamate receptor 5 precursor.
GN GRM5 OR GPRC1E OR MGLUR5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX SEQUENCE FROM N.A.
RP TISSUE-Brain;
RC TISSUE-Brain;
RA Minakami R., Katsuki F., Sugiyama H.;
RA "A variant of metabotropic glutamate receptor subtype 5: an
RT evolutionally conserved insertion with no termination codon.";
RT Biochem. Biophys. Res. Commun. 194:622-627(1993)
CC -I- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
CC MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM AND GENERATES A CALCIUM-ACTIVATED

```

```

CC CHLORIDE CURRENT.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 5A (SHOWN HERE) AND 5B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER BY AN INSERTION OF
CC 32 RESIDUES.
CC -I- TISSUE SPECIFICITY: WIDELY DISTRIBUTED IN NEURONAL CELLS OF THE
CC CENTRAL NERVOUS SYSTEM.
CC -I- MISCELLANEOUS: ACTIVATED BY QUISQUALATE > GLUTAMATE > IBOTENATE >
CC TRANS-1- AMINOCYCLOPENTYL-1,3-DICARBOXYLATE.
CC -I- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR1.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10891; NOT ANNOTATED_CDS.
CC EMBL; S64315; AB27666.1; -.
CC PIR; A42916; A42916.
CC GCRDB; GCR_0444; -.
CC GCRDB; GCR_0760; -.
CC InterPro: IPR001828; ANF_receptor.
CC Pfam; PF00003; 7tm_3; 1.
CC Pfam; PF01094; ANF_receptor; 1.
CC PRINTS; PR00248; GPCRMR.
CC PROSITE; PS00979; G-PROTEIN_RECP_F3_1; 1.
CC PROSITE; PS00980; G-PROTEIN_RECP_F3_2; 1.
CC PROSITE; PS00981; G-PROTEIN_RECP_F3_3; 1.
CC PROSITE; PS50259; G-PROTEIN_RECP_F3_4; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC MultiGene family; Alternative splicing.
CC FT SIGNAL 1 20
CC FT CHAIN 21 1203 METABOTROPIC GLUTAMATE RECEPTOR 5.
CC FT DOMAIN 22 578 EXTRACELLULAR (POTENTIAL).
CC FT TRANSSEM 579 601 I (POTENTIAL).
CC FT DOMAIN 602 615 CYTOPLASMIC (POTENTIAL).
CC FT TRANSSEM 616 636 II (POTENTIAL).
CC FT DOMAIN 637 647 EXTRACELLULAR (POTENTIAL).
CC FT TRANSSEM 648 666 III (POTENTIAL).
CC FT DOMAIN 667 692 CYTOPLASMIC (POTENTIAL).
CC FT TRANSSEM 693 713 IV (POTENTIAL).
CC FT DOMAIN 714 736 EXTRACELLULAR (POTENTIAL).
CC FT TRANSSEM 737 758 V (POTENTIAL).
CC FT DOMAIN 759 771 CYTOPLASMIC (POTENTIAL).
CC FT TRANSSEM 772 794 VI (POTENTIAL).
CC FT DOMAIN 795 800 EXTRACELLULAR (POTENTIAL).
CC FT TRANSSEM 801 826 VII (POTENTIAL).
CC FT DOMAIN 827 1203 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 377 377 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT VARSPLIC 876 907 MISSING (IN ISOFORM 5A).
CC SO SEQUENCE 1203 AA; 131885 MW; 99CA51E9E11C1E4A CRC64;

```

```

Query Match 16.28; Score 727; DB 1; Length 1203;
Best Local Similarity 27.28; Pred. No. 3.6e-45;
Matches 235; Conservative 151; Mismatches 356; Indels 122; Gaps 33;

Qy 34 LFGYVLAGLPEPLSGCQVHRPEV-TICDRSCSF--NENGYHUFQMRGVEINNST 90
Db 31 MFGIITIGALFSVH-----HGFYVDKVERKKGAAREQIGIORVEMMTLRINS DP 83
Qy 91 ALPNITLGVLYDVCSDSANVYATLRYLSLPGHHITELQGLDLYS----- 137

```

Db 84 TLLPNTTLCCEIRDSCHSA-----VALEOSIEFIRDSLSSEEEGLVRCVDGS 133
 QY 138 -----PVLAVIGPDSTRRAATALLSPFLVPMISVYASSETLSVKRQYPSFLTIPN 191
 Db 134 SSFRRKPTLVIGVIGRESSSVAILOVONLLOLFNIPOLATYATSMDSKDLFKYFMNVVS 193
 QY 192 DKYQVETVYLLQKREWTWISLVGSSDDYGQGLGVALENQATGQICIA--FKDIMPESA 249
 Db 194 DAQQAARAWMDIYKRYWMTYVSAVHFEGNAGESMEAFKMSAKEGICIAHSKYISNAGE 253
 QY 250 QYGDERMOLMHLNAGATVYVVSRROLARVFESVYVLTNLTGK--VWVASEANALSH 308
 Db 254 QSDPKLKLKLRSHLPK--ARVAVACEGKTVVGLLMMARLIGAEFLLDSDGVADRYD 311
 QY 309 IGVGPGIQRIGWLVGAIQRAVPGKAFEEAVARADKRAPPCCHGSCWSSNOLCRECO 368
 Db 312 VV--DGYQR-EAVGGITIKLOS--PDYKWPDDYLL--KLRPENLNLRPM-----FQ 355
 QY 369 AFMAHTMP-KLKAFSM-SSAYN-----AVRAYVAHGLH--QL 403
 Db 356 EFMOHRFQCRLEGEFAQENSKYNTKNSLTLRTHHYQDSKMGFVINAISYMAVGLHNMQM 415
 QY 404 LGCA--SGACSR-----GVYPMQLEQIHXYHFL-LHKDYAFNDNRDPLSYNTIAMD 455
 Db 416 SLCPGVAGLCDAKPIDGR-----KLDSIMKTNFTGVSGDMILFDENGSPGRYTEIMNEK 471
 QY 456 KNGPKWTFVLSSWSPQOLINETKIQMHGKDNQVPSVSCSDCLEGHORVY-TGFHH 514
 Db 472 EMGKQY--FLYINVGWMDNELKMDDEV-W-SKNNIINSVCSPEKQOIVIRKGEVS 528
 QY 515 CCEPCVPCGAGFFLNKSDLYRCOPCGKEWAPBEGSOTCPRIYVFLALREHTSWVLLAAN 574
 Db 529 CCMTCPCCKENEVY--FDEYTCACOLGSMPTDGLCGLIPQYLRMDPDEPIAVVFA 586
 QY 575 TLLTLLLTAGLFAMHILTPVRSAGRLCFMLGSLAAGSGLYGFGEPPRAPCLLR 634
 Db 587 CIGLATLTVYVYFIITRPTPVAKSSRELCTILAGICLYCTFCLIAKPRQIYCYLO 646
 QY 635 QALFALGFTIFLSCLTVSFOLILF-----KFTKVPPTFYHAWYONHAGLFVMISSAA 689
 Db 647 RIGIGSPMSYSALVTKNRIARILAGSKKICFKPRFMSCAQLVLAFLITICL-----702
 QY 690 QLLICTLWVYVTPLEPARYQRPPLH--VLMECTETNSLGFILAFYNGLLSISAFASY 747
 Db 703 QLGIIYALFIMEPPDLMHDY--PSIREYLLICNTFN-LGVVTPLENGLLILISCFYAF 758
 QY 748 LKQDLPENNEKCVTFSLFNFSWIAFTASVDKLYLAANMAALS-SLSGFFG 806
 Db 759 KTRNVPANNEKAYIAFWYTCIIMLAF--VPIYGSNRYKITMCFSVSLSATVALGC 815
 QY 807 YELPKCYVILCRPDNSTEHPQAS 830
 Db 816 MEVPKYYIILAKPERNVSAFTTS 839

RESULT 8
 MGR1_HUMAN STANDARD: PRT: 1194 AA.
 ID_MGR1_HUMAN
 AC_013255: 013255: 014757; 014758;
 DT_01-NOV-1997 (Rel. 35; Created)
 DT_01-NOV-1997 (Rel. 35; Last sequence update)
 DT_16-OCT-2001 (Rel. 40; Last annotation update)
 DE_Metabotropic glutamate receptor 1 precursor.
 GN_GRI1 OR GPRC1A OR MGLUR1.
 OS_Homo sapiens (Human).
 OC_Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC_Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX_NCBI_TaxId=9606;
 RN_SEQUENCE FROM N.A.
 RP_MEDLINE=96029774; PubMed=7476890;
 RA_Desai N.A., Burnett J.P., Mayne N.G., Schoepp D.D.;
 RT "Cloning and expression of a human metabotropic glutamate receptor 1

RT alpha: enhanced coupling on co-transfection with a glutamate transporter.";
 RL Mol. Pharmacol. 48:648-657(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97231349; PubMed=9076744;
 RA Stephan D., Bon C., Holzwarth J.A., Galvan M., Pruss R.M.;
 RT "Human metabotropic glutamate receptor 1: mRNA distribution, chromosome localization and functional expression of two splice variants".
 RL Neuroparmacology 35:1649-1660(1996).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM. MAY PARTICIPATE IN THE CENTRAL ACTION OF GLUTAMATE IN THE CNS, SUCH AS LONG-TERM POTENTIATION IN THE HIPPOCAMPUS AND LONG-TERM DEPRESSION IN THE CEREBELLUM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS. STRONGEST, TO MGLUR5.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U31215; AAA87843.1; -;
 DR EMBL: U31216; AAA87844.1; -;
 DR EMBL: L76627; AAB05337.1; -;
 DR EMBL: L76631; AAB05338.1; -;
 DR GCRDB: GCR_1825; -;
 DR GCRDB: GCR_1826; -;
 DR GCRDB: GCR_1982; -;
 DR GCRDB: GCR_1983; -;
 DR MIM: 604473; -;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR00337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_REC_P3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_REC_P3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_REC_P3_3; 1.
 DR PROSITE: PS00259; G_PROTEIN_REC_P3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Multigene family; Alternative splicing.
 FT SIGNAL 1 18
 FT CHAIN 19 1194
 FT DOMAIN 19 592
 FT TRANSMEM 593 615
 FT TRANSMEM 616 629
 FT TRANSMEM 630 650
 FT TRANSMEM 651 661
 FT TRANSMEM 662 680
 FT TRANSMEM 681 706
 FT TRANSMEM 707 727
 FT TRANSMEM 728 750
 FT TRANSMEM 751 772
 FT TRANSMEM 773 785
 FT TRANSMEM 786 808
 FT TRANSMEM 809 814
 FT TRANSMEM 815 840
 FT DOMAIN 841 1194
 FT DOMAIN 1014 1035
 FT DOMAIN 1067 1081
 FT DOMAIN 1095 1130
 FT DOMAIN 1142 1194
 FT CARBOHD 98 98
 FT CARBOHD 223 223

POTENTIAL.
 METABOTROPIC GLUTAMATE RECEPTOR 1.
 EXTRACELLULAR (POTENTIAL).
 I (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 II (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 III (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 IV (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 V (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 VI (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 VII (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 GLN/PRO-RICH.
 GLN/PRO-RICH.
 ASP/GLU-RICH (ACIDIC).
 SER-RICH.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match	16.0%	Score 719;	DB 1;	Length 1194;
Best Local Similarity	26.2%	Pred. No. 1.4e-44;		
Matches 225;	Conservative 158;	Mismatches 368;	Indels 108;	Gaps 307;

QY	34	LPBGYLLAGLEPLHSGCLOVHREPV--TLDCRSCS--FNHGHYLFQAMRLAGEEINNST	90
Dd	41	MDGVOIIGALEFSVH-----HQPAPKAPVERKCGEIREQYQIRVAMHHTIDKINADP	93
QY	91	ALPNTLITGLQIYVCSDSANVTATLRYLSIPGCHITELGGDL-----	134
Dd	94	VLPNTLITGSEIRDSCHWS-----SVALEQSIIEFTIRDSLISIRDEKDGINRLPD	143
QY	135	-----HYSPTLVAIVGPSTNRAATTAALLSPFLVPMISVYASASSTLSVKRQYPSER	187
Dd	144	GOSLPRGCKTKRPAGVIGRPGSSVAIOYONILDLFDIPQIYASTSITDLSDKTLKYFLR	203
QY	188	TIPNDKYQVETMVLLOKFCGWTWISLVGSSDDYQGLQVQALENATGOGICIFKDIMPF	247
Dd	204	VVPSDTLOARAMDIVIKRYNTYVSAVHTBEQNGESGMDAFKELAAEGICIAHSD--KI	261
QY	248	SAOYQDEMOCOLMHILAO--AGATVVVVYFSSROLARPFESVYLITNLTKV--WASAPMA	304
Dd	262	YSNAGESSFDRLKRLKREBRYPKARVVVCFEGMTVRLLSAMRLGVAGFEFSLSGDMA	321
QY	305	LSRHITGPGIORIGMTVLGVALIKRAVPGKAFPEAV--ARADKAPRPHKSGWSCSNQ	362
Dd	322	DRDEYIEBEYEVANG-----GITIKLOS--PEVRSFDYILKLRLDNTNPMFPREWMQRFQ	377
QY	363	LCRECOAFMAHTM--PKIKAF-----SMSAY-----NAYRAVVAVANGL-----HOLL	404
Dd	378	--CR-----LPGHLEPNPFKRICTGENSLEENYVODSKMGFVAINAYAMAHGQMHALC	432
QY	405	GCASGACSGRG--VYPMQDLBEDIHNVHFL--LHKQVYAANDNRDLSTNIIANMOM--NOPKM	461
Dd	433	PGHVGLCDAMPRIDGSKLDLRLSSFTIGVSGEEVWFDEKGDAPGRVDIMNLOTTEANRY	492
QY	462	TFTVLGSSWTSPVOLINETKTIQMHGKDNOPKXSSCSDCLEGHORVY--TGFIHSCPECV	520
Dd	493	DYVHNG--THMEGLVINTDYKIQNN--KSGVVRBSVCEPCKQKQIKYIRGGEVSCMIGT	548
QY	531	PCGACTPLANKSDLYRCOPCKGEENAPRECSOTCFPRYVFTALREHTSWILLANTLILL	560
Dd	549	ACKENEYV--ODEPTCKACDGMWPMNADLTGCEBPIPRVYLEWSNIEPIIIAIESCLGIV	606
QY	581	LLGTAGLEAMHLDIPRVYRSAGRGCFMLTGLSLAGSGSIVGFECEPRPACILROALFAL	640
Dd	607	TLFPTLITLVLYRDPVYKSSRELCTIILAGIFLGIYVCFPIILAKPPTTSCYIORLLVGL	666
QY	641	GFTIFLSCLTVRSPQLIIIF-----KSTKYPRFYHAMVONHGAGLFVMTIS--AAOL	692
Dd	667	SSAMQYSALVTKRIARILLAGSKKIKICTRPRMSAMAO-----VIIASLISVOLT	719
QY	693	ICLTMVVTWTFLPAREQRPFLVMELECTEPLNSGLFLAPLYNGLSISAFACSYLEKDL	752
Dd	720	LVVTLILMEPMPILSYPSIKEVLI--CMTNSNGVAPAPLQNGLLIMCSTYVAFKFRNY	777
QY	753	PENTNEACVFEFSLLENFVSIAPFTTASYVDG--KYLPAANMMAAGSLSSGFGGYLPK	811
Dd	778	PANREAKYIAPFTYTTCIIMLAF-----VPIYFGSNYKIITTCRAVLSVTVALGCMFTPK	834
QY	812	CYVILCRPDLNSTEHPQAS	830

DB		835	MIIIAKPERNRSAFTS	853
	RESULT	9		
ID	MGR4_RAT	STANDARD:	PRT:	912 AA.
AC	P31423.			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Metabotropic glutamate receptor 4 precursor.			
CN	GRIA4 OR GPRC1D OR MGLUR4.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCB1_Taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain.			
RX	MEDLINE=92110002; PubMed=1309649;			
RA	Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;			
RL	"A family of metabotropic glutamate receptors.";			
RL	Neuron 8:169-179(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain.			
RX	MEDLINE=93332699; PubMed=8338667;			
RA	O'Hara P.J., Sheppard P.O., Thoenes H., Venezia D., Haldeman B.A.,			
RA	McGrath V., Houmed K.M., Thomson C., Gilbert T.L., Mulvihill E.R.;			
RT	"The ligand-binding domain in metabotropic glutamate receptors is			
RT	related to bacterial periplasmic binding proteins.";			
RL	Neuron 11:41-52(1993).			
CC	- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR			
CC	IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE			
CC	ACTIVITY.			
CC	- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS. PREDOMINANT			
CC	EXPRESSION IS SEEN IN THE GRANULE CELLS OF THE CEREBELLUM.			
CC	- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.			
CC	STRONGEST, TO MGLUR6.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL, M92077; ; NOT_ANNOTATED_CDS.			
DR	EMBL, M90518; AAA93190.1; ;			
DR	PIR, JH0563; JH0563.			
DR	GCRdb, GCR_0352; -;			
DR	GCRdb, GCR_0363; -;			
DR	InterPro, IPR001828; ANF_receptor.			
DR	InterPro, IPR000337; GPCR_Mgr.			
DR	Pfam: PF00003; 7tm_3, 1.			
DR	Pfam: PF01094; ANF_receptor; 1.			
DR	PRINTS; PR00248; GPCRMGR.			
DR	PROSITE; PS00979; G_PROTEIN_RECPE_F3.1; 1.			
DR	PROSITE; PS00980; G_PROTEIN_RECPE_F3.2; 1.			
DR	PROSITE; PS00981; G_PROTEIN_RECPE_F3.3; 1.			
DR	PROSITE; PS0259; G_PROTEIN_RECPE_F3.4; 1.			
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;			
KW	Multiogene family.			
FT	SIGNAL	1	32	POTENTIAL.
FT	CHAIN	33	912	METABOTROPIC GLUTAMATE RECEPTOR 4.
FT	DOMAIN	33	587	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	588	610	I (POTENTIAL).
FT	DOMAIN	611	624	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	625	645	II (POTENTIAL).
FT	DOMAIN	646	656	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	657	675	III (POTENTIAL).

DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR PRINTS: PRO0248; ANF_receptor; 1.
 DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS00259; G_PROTEIN_RECP_F3_4; 1.
 DR G-protein coupled receptor; transmembrane; Glycoprotein; Signal;
 KM Multigene family.
 FT SIGNAL 1 32
 FT CHAIN 33 912
 FT DOMAIN 33 912 METABOTROPIC GLUTAMATE RECEPTOR 4.
 FT TRANSMEM 588 610 I (POTENTIAL).
 FT DOMAIN 611 624 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 625 645 II (POTENTIAL).
 FT DOMAIN 646 656 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 657 675 III (POTENTIAL).
 FT DOMAIN 676 699 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 700 720 IV (POTENTIAL).
 FT DOMAIN 721 750 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 751 772 V (POTENTIAL).
 FT DOMAIN 773 785 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 786 808 VI (POTENTIAL).
 FT DOMAIN 809 821 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 822 847 VII (POTENTIAL).
 FT DOMAIN 848 912 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 484 484 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 912 AA; 101867 MW; 4A2F36B53A2EA5A CRC64;

Query Match 15.7%; Score 704.5; DB 1; Length 912;
 Best Local Similarity 26.9%; Pred. No. 1.1e-43;
 Matches 228; Conservative 142; Mismatches 379; Indels 99; Gaps 31;

QY 34 LPGYLLAGLFPRLHSGCLQVHRREYVLCDRSCSFNENGYHLPQAMRLGVEINNNSTLL 93
 DB 45 IDGGITLGLFPPVGRSGEPCGELK-----KEGGIHRLEAMFALDRINNDPDL 96
 QY 94 PNILGQLVDCSDSANVYATLRLVSLPGHTELOG-----DLHYSPTYLAV 143
 DB 97 PNILGARIDTC--SRDTHLEBSLTFV-QALLEKGTETVRCSGGPPIITTKERYVGV 153
 QY 144 IGPSTNRATTAALLSPFLVPMISVYASSETLSVKROYPSFLRTIPNDKYQVETWVLL 203
 DB 154 IGAGSSSVSIVANILRLFKIPQISVASTADPLSDNSRYDFSRVPSDVTQAQAMVDIV 213
 QY 204 QKEFWTISLVGSSDDYGGQGVQA-LENOATGOGICIAFKRIMPESAQVDEBRQCLMRH 262
 DB 214 RALKMNVSVYASGSGYGESEVFAIKSRDEGVCAISVYKIPREPKAG-EFKILIRRL 272
 QY 263 LAQAGATVVVVFSSRQARVFESESVLTNLGK-VWVAASEAMALSRHTGTGPGIQRIGMV 321
 DB 273 LEISNAVAIIFANEDIRYLEARRANOTGHEFWMSDSW--GSKIAPVLIHEVEAG 330
 QY 322 LGVAIOKRAVGLKAFEEAVARADKEARPCCHKGSWCS---SNOLRECOFAFAMTPRK 377
 DB 331 AVTLIPKRM--SVRGDFRFFSSRLDNNR--RNIMWFEFEDNPHCK--LSRHALKK 381
 QY 378 -----LKAFSMSSAYN-----AYRAVIVAHGLH---QLLGASGACSR-GRV 416
 DB 382 GSHVKKCTNRERRIGQDSAYEEOGEKVQVPIADVAMGHALMHARDLCPGRGLCPRRMPV 441
 QY 417 YPMOLLEOIRKVFPL-LHKDTVAFNDRNDPLSNIITAMDNNGKMPFTVSGSTWSPVQ 475
 DB 442 DGTOLLEKIRRVNPSGIAGNVTYTNENGDAGRIDYIYQLRNDSAEKTVGS--WTD-H 498
 QY 476 LINETKIQWGHKDNQVPKSVCSDDCLEGHORVYTGPHHCCEGVPCGAGTFLNKSDLIR 535
 DB 476 LNIETKIQWGHKDNQVPKSVCSDDCLEGHORVYTGPHHCCEGVPCGAGTFLNKSDLIR 535

DB 499 LHLRIERMHPMGSGOQLPRISICSLPCQGERKKTYKMGPCMHCEPCTG--YQVQVDRY 556
 QY 536 COPCKEEMAEDEGSCOTCPPRVVFALAREHTSWILLANTLLLLCTAG-----LFA 589
 DB 557 CKTCPYDMRPENRRTG--RPPIIKLEWGSFVWLP-----LFAVGIATLFEVITFV 610
 QY 590 WHLDPPVRSAGRLCEFLMGLSAAAGSLGFGCEPRPACLRQALFALGFTFLSLC 649
 DB 611 RYNDPIYKAGRELSTYLLGIFLCYATTFMLAEPLGTCSLRLRFLGKMSIYVAL 670
 QY 650 TVRSFOLIIF--KFSTKVPFFYHAWQNHAGLFVMISSAOLLICLTWLVVWTPDPA 706
 DB 671 LTKNRIYRIREQKRSVSADRFISPAQL--ATFFSLIS--LQLLGICVWFVDPDSHV 726
 QY 707 REYQ-----RPHLVMLECTETNSLGFILAFYLLGLISAFACSTLGKDPENYEA 759
 DB 727 VDFODORTLDRPRFAGV-LKC-DISDSLILCLLGLSMILMTCTVYAAKTGVPETFNEA 784
 QY 760 KCVTFSLLENFVSNIA-----FETTASVYDGKYLPAANMAGIS--SLSSGFGYFLPKCYV 814
 DB 785 KPFGFTMTTCTIWLAFIPFIFGTISQADKLYIQTTITLVSVLSASVSLGMLYKPYI 844
 QY 815 ILCRPDLN 822
 DB 845 ILFHEQN 852

RESULT 11
 ID MGR2_RAT STANDARD; PRT; 872 AA.
 AC P31421;
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 40, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 2 precursor.
 GN GRM2 OR GPCRB OR MGLUR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92110002; PubMed=1309649;
 RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.,
 RL Neuron 8:169-179(1992).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 CC MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
 CC MAY MEDIANE SUPPRESSION OF NEUROTRANSMISSION OR MAY BE INVOLVED IN
 CC SYNAPTOGENESIS OR SYNAPTIC STABILIZATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS AND PROMINENT
 CC EXPRESSION IS SEEN IN GOLGI CELLS OF THE CEREBELLUM AND SOME
 CC PARTICULAR NEURONAL CELLS IN OTHER BRAIN REGIONS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; M92075; ; NOT_ANNOTATED_CDS.
 CC PIR; JH0561; JH0561.
 CC HSP; P06612; 1ECL.
 CC GCRDB; GCR_0361; .
 CC InterPro: IPR001828; ANF_receptor.
 CC InterPro: IPR000337; GPCR_Mgr.
 CC Pfam; PF00003; 7tm_3; 1.

DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PS00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECPEP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECPEP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECPEP_F3_3; 1.
 DR PROSITE; PS0259; G_PROTEIN_RECPEP_F3_4; 1.
 DR G-protein coupled receptor; transmembrane; Glycoprotein; Signal;
 MW Multigene family.
 FT SIGNAL 1 18
 FT CHAIN 19 872
 FT DOMAIN 19 567
 FT TRANSMEM 568 590
 FT DOMAIN 591 604
 FT TRANSMEM 605 625
 FT DOMAIN 626 636
 FT TRANSMEM 637 655
 FT DOMAIN 656 679
 FT TRANSMEM 680 700
 FT DOMAIN 701 725
 FT TRANSMEM 726 747
 FT DOMAIN 748 760
 FT TRANSMEM 761 783
 FT DOMAIN 784 793
 FT TRANSMEM 794 819
 FT DOMAIN 820 872
 FT CARBOHYD 203 203
 FT CARBOHYD 286 286
 FT CARBOHYD 338 338
 FT CARBOHYD 402 402
 FT CARBOHYD 547 547
 SEQUENCE 872 AA; 95773 MW; 1E/4CABD6AD4BED9 CRC64;

Query Match 15.4%; Score 692.5; DB 1; Length 872;
 Best Local Similarity 27.5%; Pred. No. 7.9e-43;
 Matches 245; Conservative 121; Mismatches 361; Indels 165; Gaps 38;

QY 33 TLPGDYLAGLFLPHSGGLQVNRPEVTLCDSGSENEH-GYHLFQAKRLGVEEINNSTA 91
 DB 27 TLEGDLVGLGFLPVHQ-----KGGP-----AEECGPVNEHRIORLEAMFALDRINRDPH 77
 QY 92 LLPRTTGLYOLVDCSDAN-----VYATRLVLSLPGCHHIELOQDLHYSS--PT-V 140
 DB 78 LLPGRLAHLLIDSCSKTHALEQALDFVRLASL-RGADGSHNICPDGSDYATHSDAPYAV 136
 QY 141 LAVIPDSTNRATYTAALSPFLVPMISYAASETLSVKRQYPSFLRTIPNDKYOVETVY 200
 DB 137 TGVIGGSYSDVSIOYANLRLFOIPISTASAKLSKSRDYRARTYPRPFQAKAMA 196
 QY 201 LLLQKFGWTWISLVGSSDDYQGLQVQALENQATGGGICIAEKDIMPFSAQVGDERMOCIM 260
 DB 197 ELIRFNNYTVSTVASEDYGEGTIEAELELARANICVA-----TSEKVRAMSRAAF 250
 QY 261 RHLQA-----GATVVVSSRQLARVFEESVYVLTNLGKVVAVASEANLSHITGVPI 315
 DB 251 EGVVALLQKRSARVAVLTSEDAKRELLAATQRLN-ASFTWVAADGWALESV----- 303
 QY 316 QIRGAVLGLVATOKRAVPGLKAFEEA-YARAD-----KEAPRCPCHKSM-----C 358
 DB 304 -----VAGERAAEGATITELASYPISDFASYFQSLDPNNRNPNRFRFEMERFHC 355
 QY 359 SNNQLCRCQAFMAHTMKLAFSSMSAYNAYRAYVAHGHOLLG-----CAGS 409
 DB 356 SFRQ--RDCAAHSLRAVFEQESKIMFVYN--AVYAMAHALHNNHRAALCPMTTHLCDAM 410
 QY 410 ACSRGVYPMQLLEQIHHVHLLH-----KOTVAAANDNRDPLSSNITIAMDMN 457
 DB 411 RPNVR-----RLYK-DEVLNKKFDPAPRPADTDEVDREDRGIGRNNIFTYLA 461
 QY 458 GP-KMTFTVLGSSWSPOLNINETKIOHMKD-NOVPKSSVSSCLLEGHORVYTGFIHC 515
 DB 462 GSGRRYKQVGV--YVAE-GLITLDTSFITWASAPRLPASRCSFECLONEKVSYPGEVC 518

OY 516 CREVPCGAGTFLNKSIDLVCOPCKEMAEPCGOT-CPPTVFLALREHTSWL--LA 572
 DB 519 CWLICPCPYEX--RLDEFTCADCGLGW-PNASITGCFELPOEY--IRMGAMAVGCVT 573
 OY 573 ANTLILLILLITAGLIPAHMLDPPVYRSAGGRICELMIGSLAAGSSSLVGFEEPRPACL 632
 DB 574 IACIALATFLVGLVFNRAVNPVYKASGRELCYLLGGVFLCYCMTFVFLAKPSTAVCT 633
 OY 633 LRQALFALGFETFLSCLVVRSPOLITIF--KFTKVPFTFYHAMVQNHGAGLEWISSNA 689
 DB 634 LRLRLGIAFSVCSYCALTKTRIRITICGAREGQORPF-----ISPAS 678
 OY 690 QLLICL-----TWLVVTP-----LPAREYRPHLVMLCT--ETNSLGFIL 730
 DB 679 QVAICLALISGOLLIVAAVLVEAGTGETPERR-----EVTILRCNHRDASMLG--- 730
 OY 731 AFLVGLGISAFASYLGKDPENYNEAKCTSEFLLENFVSWIAPFTTASYDOKYLP 790
 DB 731 SLAYVLLIALCTLYAFTRKCPENFNEKFTGFTMYTTCILMLFLPIFYTSSDYRQ 790
 OY 791 ANMMAGLSLSSG--FGGYFLPKCVIICRPDLNSTEHFOASIDYTRCGS 840
 DB 791 TTTCVSVSLSSGVVLCGLFAPKHLITILFQPKNVVSH-----RAPTRFGS 837

RESULT 12
 MGR3_RAT
 ID MGR3_RAT STANDARD: PRT: 879 AA.
 AC P31422:
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 3 precursor.
 GN GRM3 OR GPRC1C OR MGLUR3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92110002; PubMed=1309649;
 RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RT "A family of metabotropic glutamate receptors."
 RL Neuron 8:169-179(1992).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS. PREDOMINANT
 EXPRESSION IS SEEN IN THE NEURONAL CELLS OF THE CEREBRAL CORTEX,
 DENTATE GYRUS, AND GLIAL CELLS THROUGHOUT BRAIN REGIONS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONEST. TO MGLUR2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC EMBL; M92076; -; NOT_ANNOTATED_CDS.
 DR PIR: JH0562; JH0562.
 DR GCRB; GCR_0362; -.
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR Pfam; PF01094; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PS00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECPEP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECPEP_F3_2; 1.

FT TRANSMEM 575 597 I (POTENTIAL).
 FT DOMAIN 598 611 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 612 632 II (POTENTIAL).
 FT DOMAIN 633 643 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 644 662 III (POTENTIAL).
 FT DOMAIN 663 686 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 687 707 IV (POTENTIAL).
 FT DOMAIN 708 732 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 733 754 V (POTENTIAL).
 FT DOMAIN 755 767 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 768 790 VI (POTENTIAL).
 FT DOMAIN 791 800 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 801 826 VII (POTENTIAL).
 FT DOMAIN 827 877 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 877 AA; 98619 MW; 66F28663CE35F740 CRC64;

Query Match 15.2%; Score 684; DB 1; Length 877;
 Best Local Similarity 25.7%; Pred. No. 3.3e-42;

Matches 225; Conservative 140; Mismatches 361; Indels 150; Gaps 33;

DB 30 EIKIGEDLVGLGLPINE-----KGTGEEGRINEDGRIQLEMLFAIDENKD 80
 QY 31 DFTLPDYLVLGLPLHSGCLQVRRHREVTLLCDRSCFNE-HGYHLFOAMRLGVEEINNS 89
 DB 90 TALLNITIGYLDYVCSANVATRLSL-----PGCHHIELQGL 133
 DB 81 DYLLEGVKLVGHILPTC--SRDTYALDSELEFVRASLTKVDAEYKCPDGSYAIQENIP 138
 QY 134 LHYSPFLVAVIGPDSNTRATTAALSPFLVAVISYASSETLSYKQPSFLRTIPNDK 193
 DB 139 L-----IAGVIGSYSVSIQVANNLRLEFQIPIQISASLSKSRDYARTYPPPF 193
 QY 194 YQVETMVLLOKFGMTWISLVGSSDDYGOLOVALENOATGOGICIAFDIMPFAQVD 253
 DB 194 YQAKMAELREFNMTYVSTVASEGYGETGIEAFEOEARLNICATAE-----KVGR 247
 QY 254 EMO-----CLMHLAO-AGATVAVVSSNQLARVFESVVLNINLQKVVAVASEAMLSH 308
 DB 248 SNIRSYSDYIELLOKPNARVAVLEPMSDDSRBELIAASRAN-ASFTVAVDGMGQDS 306
 QY 309 ITGVPGIORIGVAVLOKRAVPGIKAFEEAYARADKEAPPCHKGSV-----C 358
 DB 307 I--IKGSEHV--AYGATLELASQVRODRYFQSLN--PYNHNNPWFDFWQKPGC 359
 QY 359 S---SNQCRBEOATMAHMPKLAFSMSSAYN-----AVRAYAVAHGLHQ---LL 404
 DB 360 SLONKRNHRVCDKHLAID-----SSNYEOKSKIMEFVNAVAYAMAHALHKMORTJC 410
 QY 405 GCASGACSGRGVYPMQLLEQIHVFLHKDVAFNNDNDP-----LSYNI 452
 DB 411 PWTYLCIDAMKILDG---KKLYKDYLLKINFTAPFNPKDADSIYKFTFGGOMGRTYNF 467
 QY 453 AWDWNGPKWTFTVLGSSSTWSPVQNLINETKIQMHGKDNQVPSVSSDCLGSHORVVTGF 512
 DB 468 NFQVNGKSYLYKVGH--WAET-LSLDVNSIHW--SRNSVPTSQSDPCAPAPNKKMOMG 522
 QY 513 HHCCEVPCGAGTFLNKSNDLYRCOPCKEENAPESQTCFRTYVFLAREHTS- 568
 DB 523 DVCWCICIPCPEYEYL--ADEFTCMCGSGOWPTADLTCTCYLDPEDY--IRWEDAWAIGP 578
 QY 569 VLLANTLILLLLTGAGIFAMHLDTPVYRAGRGICFLMGLSLAAGSSSLGFFCEPR 628
 DB 579 VTIACLGCTCMVVT--VFIAHNTPIVAKASGRELCYLLGVLGYSTCMTEFFIAKPS 636
 QY 629 PACLRQALFALGFTIFSLCTVRSFOLLIF--KSTKVPYTFYHAWQNHGAGLFWMI 685
 DB 637 VICALRRIGLSSFAICYSALLTKINCIAIRFDGVKNGAKRKF-----I 681

QY 686 SSAOLLICITLVV-----VWTLPAERYQRP-----HVMIECTETNSLGLTAPL 733
 DB 682 SPSSQVETICGLIVQVWVWLLLEAPGTRRYTLAEKRETVILKCNKDS-SMISLT 740
 QY 734 YNGLSISFACSYLCKDIPENYNEAKCYTESLLENFVSWINAFPTTASYDCKYLPAAAM 793
 DB 741 YDVIVILCTVYAFATRKCPENENKCFGTMYTTCIIMAFLEPIYVTSDDYVQTT 800
 QY 794 MAGLSLSSGF---GGYFLPKCYVILCRPDLNSTEH 826
 DB 801 MC-IVSLSGFVVLGLFAPKVIHLLFPQKNVYH 835

RESULT 14
 MGR2_HUMAN STANDARD; PRT; 872 AA.
 ID MGR2_HUMAN
 AC 01416;
 DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Metabotropic glutamate receptor 2 precursor.
 GN GRM2 OR GPRC1B OR MGLUR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95346007; PubMed=7620613;
 RA Flor P.J., Lindauer K., Putner I., Ruegg D., Lukic S., Knopfel T.,
 RA Kuhn R.;
 RT "Molecular cloning, functional expression and pharmacological
 RT characterization of the human metabotropic glutamate receptor type
 RT 2";
 RL Eur. J. Neurosci. 7:622-629(1995).
 CC -I- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 CC MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
 CC MAY MEDIANE SUPPRESSION OF NEUROTRANSMISSION OR MAY BE INVOLVED IN
 CC SYNAPTOGENESIS OR SYNAPTIC STABILIZATION.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- TISSUE SPECIFICITY: WIDELY EXPRESSED IN DIFFERENT REGIONS OF THE
 CC ADULT BRAIN AS WELL AS IN FETAL BRAIN.
 CC -I- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC -I- STRONGEST, TO MGLUR3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L35318; AAA76855.1; .
 CC GCRDB: GCR.1846; .
 DR DR MIM: 604093; .
 DR DR InterPro: IPR001828; ANF_receptor.
 DR DR InterPro: IPR000337; GPCR_Mgr.
 DR DR Pfam: PF00003; 7tm_3; 1.
 DR DR Pfam: PF01094; ANF_receptor; 1.
 DR DR PRINTS: PR00248; GPCRMBR.
 DR DR PROSITE: PS00979; G_PROTEIN_RECPE_F3_1; 1.
 DR DR PROSITE: PS00980; G_PROTEIN_RECPE_F3_2; 1.
 DR DR PROSITE: PS00981; G_PROTEIN_RECPE_F3_3; 1.
 DR DR PROSITE: PS00982; G_PROTEIN_RECPE_F3_4; 1.
 DR DR G-protein coupled receptor; Transmembrane; glycoprotein; signal;
 DR G-protein family.
 DR MULTIGENE FAMILY.
 FT SIGNAL 1 18
 FT CHAIN 19 872 METABOTROPIC GLUTAMATE RECEPTOR 2.
 FT DOMAIN 19 567 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 568 590 I (POTENTIAL).
 FT DOMAIN 591 604 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 605 625 II (POTENTIAL).

FT DOMAIN 626 636 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 637 655 III (POTENTIAL).
 FT DOMAIN 656 679 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 679 700 IV (POTENTIAL).
 FT DOMAIN 701 725 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 726 747 V (POTENTIAL).
 FT DOMAIN 748 760 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 761 783 VI (POTENTIAL).
 FT DOMAIN 784 793 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 794 819 VII (POTENTIAL).
 FT DOMAIN 820 872 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 872 AA: 95507 MW: 05860835C701E9D CRC64;

Query Match 14.8% Score 666.5; DB 1; Length 872;
 Best Local Similarity 27.4%; Pred. No. 6.3e-41;
 Matches 242; Conservative 126; Mismatches 366; Indels 147; Gaps 41;

QY 33 TLPGDYLAGLPLHSGLQVRRPEVTLCDRSCS-FNEH-GYHLFQAMRLGVEEINNST 90
 DB 27 TLEGDVLVGLGFVHQ-----KGGPA-----EDCGPVNEHGIORLEAMLFALDRINDP 76
 QY 91 ALLENITLGYLYVCSDSAN-----VYATLRYLSLPCOHITELGDLIHT--SPT- 139
 DB 77 HLLGVRRGAHILDCSCTHRLDQALDFVRASLS-RCGADSRSRHCPCPGSVATHDADPTA 135
 QY 140 VLAVIGPSTRRAATYTAALLSPFLVPMISYVASSSETLSVKROYPSFLTINDKYQVETM 199
 DB 136 ITGYIGGSYSVSQVAMNLFRLFOIPQISTASTAKISDKSRIDYFANTVPDPFOAKAM 195
 QY 200 VLLQKGTWISLVGSSDDYGOLGVOALENQATGCGICAFKDIIMPESAGVDERMQL 259
 DB 196 AEILRFNMVYVSTSEASGDYGETIEAFELFARNICVA-----TSEKVGAMSRSA 249
 QY 260 MRHLAQA-----GATVVYVSSROLARFEESVLTNLTKGVYVASEM-ALSRHITVVP 313
 DB 250 FEGVRRALLOKPSARVAVLFTFRSEDARELLAASORLN-ASPTWVASDGMGLSEVVASGE 308
 QY 314 GIORIGWLVGAIQKRAVPGIKAFEEAVARAD--KEAPRPHKQSW-----CSSNOLCRE 366
 DB 309 GAAL-----GATITELASP-ISDRASYFQSLDPMNNSNPFREFWEDRFCSFQO--RD 361
 QY 367 QCAFMAHMPKLKAFSSSAYNAVAYAVAGHLHQLGSCASGACSRGVYPMOLLEQIH 426
 DB 362 CAASLSRAVPPEQESKIMFYVN---AYAMAHALHNMHRALCPMTTR-----LCDAMR 411
 QY 427 KVH-FLLHKD---TVARNDRDPLSSYITIAMDNKPK-----WTFVLLGSSST----- 470
 DB 412 PVNRRRLYKDKDLNVKCDAPPRPADTHNEVFRDGGDIGNYJFTYLRASGRARYKV 471
 QY 471 --NSPVOLINETKTIOMHGXKD-NOVPKSVSCSDCLEGHORVVTGFHHCCEFCVPGAGTF 527
 DB 472 GYMAF-GITLDTSLIPMAFSPAGPLASRCSEPCIONKVKVQGEVYCWMLCIRQPEY 530
 QY 528 LMSKDYLCOPCKEEMAPESQF--CEPRVVFALREHTSWLT--LAANTLILLLLLT 584
 DB 531 --RIDEFACACGGLGYM-PNASTLGCEFLPOEY--IRMGDAWAVGPVTIACIGALATLFLV 585
 QY 585 AGFAMHLDPRVVSAGRCFCMLGSLAAGSSLYGFFGPTRPACILRQALFALGTI 644
 DB 586 LGVFNHATVVRKASGEELCYILLGVFLCYMTFETIAPRSTAVCTLRRLGLGTAFSV 645
 QY 645 FLSCITVRSFOLITIF--KFTSVPTPEYHAMVONHAGLFLVMISSAQLICT----- 695
 DB 646 CYSALLTKNTNIAHIFGARAGCAQPRF-----ISPAQVACIALISGQ 690
 QY 696 -----TWLVVTP-----LPAREYQRPPLVMECT--ETNSLGFILAFYNGILSISA 742

DB 691 ILIVVAMLVENAPGTGKETAPER-----EVLTRCNHRDASMLG---SLAYVLLIALC 742
 QY 743 FACSVLGDLPENYKACVPEFSLFNFWSTIA-----EFTTASVDGKYLPAANMAGIS 798
 DB 743 TLVANTRKCPENENAFIGFTMTTCITWALLPIFYVSSD---RVQTTMCVSVS 799
 QY 799 -SLSSFGGYFLPKCYVILCRPDLNSTEPQASIDYTRGGS 840
 DB 800 LSGSVVLGCLFAPKHLIILFQPKNVSH-----RAPTSRGS 837
 RESULT 15
 MGR6_HUMAN
 ID MGR6_HUMAN STANDARD; PRT; 877 AA.
 AC O15303;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 6 precursor.
 GN GRM6 OR GPRC1F OR MGLUR6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97358610; PubMed=9215706;
 RA Hashimoto T., Inazawa J., Okamoto N., Tagawa Y., Bessho Y., Honda Y.,
 RA Nakanishi S.;
 RT "The whole nucleotide sequence and chromosomal localization of the
 RT gene for human metabotropic glutamate receptor subtype 6.";
 RL Eur. J. Neurosci. 9:1226-1235(1997).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR4.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; U82083; AAB82068.1; -.
 DR GCRDB; GCR_2607; -.
 DR MIM; 604096; -.
 DR InterPro; IPR001828; ANP_receptor.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE; PS50259; G_PROTEIN_RECP_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multi-gene family; Vision.
 FT SIGNAL 1 24
 FT CHAIN 25 877
 FT DOMAIN 25 585
 FT TRANSSEM 586 608
 FT DOMAIN 609 622
 FT TRANSSEM 623 643
 FT DOMAIN 644 654
 FT TRANSSEM 655 673
 FT DOMAIN 674 697
 FT TRANSSEM 698 718
 FT DOMAIN 719 748
 FT TRANSSEM 749 770
 FT
 FT METABOTROPIC GLUTAMATE RECEPTOR 6.
 FT EXTRACELLULAR (POTENTIAL).
 FT I (POTENTIAL).
 FT CYTOPLASMIC (POTENTIAL).
 FT II (POTENTIAL).
 FT EXTRACELLULAR (POTENTIAL).
 FT III (POTENTIAL).
 FT CYTOPLASMIC (POTENTIAL).
 FT IV (POTENTIAL).
 FT EXTRACELLULAR (POTENTIAL).
 FT V (POTENTIAL).

FT DOMAIN 771 783 CYTOPLASMIC (POTENTIAL).
FT TRAMSMEM 784 806 VI (POTENTIAL).
FT DOMAIN 807 819 EXTRACELLULAR (POTENTIAL).
FT TRAMSMEM 820 845 VII (POTENTIAL).
FT DOMAIN 846 877 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 479 479 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 877 AA: 95436 MW: AECDF91EBDA5994F CRC64;

Query Match 14.8%; Score 666.5; DB 1; Length 877;
Best Local Similarity 25.7%; Pred. No. 6.3e-41;
Matches 232; Conservative 139; Mismatches 377; Indels 155; Gaps 33;

QY 11 LQLLISCCWAFACHSTRESSPDLFLPDDYLLAGLPLHSGCLQVRRHREVTLCDRSCS--F 68
DB 12 LVALLPLAWLAQGLARAGSVRLAGGLTLGLFVPHAR-----GAGGRACGPLK 61
QY 69 NEHGYHLFQAMRLGYEINNSTALLPNTILGYQLYDVCSDSANVATRLVLSLPGQHIE 128
DB 62 KEGGYHREAMLYALDRVNADELLPGVRLGARLLDTC--SRDTYALBQALSFV-QALIR 118
QY 129 LOGD-----LHYSB--TVLAVIGPDSTNRATTAALLSPFLVPMISTYASS 173
DB 119 GRGDGDEGVRCRPGGVPLRPAPERVAVVAGSASVIMAVNLRLFAIPQISYASTA 178
QY 174 ETLISVKROYPSFLRTIPNDKYOVETVLLLOKFGWTWISLVGSSDDYGOLGYOA-LENOA 232
DB 179 PELSDSTRIDFSSRVVPPSTYQAAWVDIVRALGMNTVSTLASSEGNIGSGVEAFVQISR 238
QY 233 TGOGICIAFKDIMPESAOYDERMOCIMHQAQATVVVVFSSROLARVFESVVLTNL 292
DB 239 EAGVCIQAQSIKIPREPCKG-EFSKYIRLMEPTNARGIITFANEDDIRVLEAARQANL 297
QY 293 TGR-VWVASEANALSHHTGVEIQIRIGVYLAIOKRAVPGIKAPEEAVARADKAPRP 351
DB 298 TGHFLMWGSDSWGAK--TSPILSLDEVAAGAITILPKRA--SIDGDOYFMRSLNNR- 352
QY 352 CHKGSW-----CSSNQLCREQAQAFMAHTMPRLKAFMSASVN----- 388
DB 353 --RNIWFAEFMEENENCKLITSSGTOSSDSTRKCTG-----EERIGDSTIEDGKVO 402
QY 389 -AYRAYVAHGLHQ-----LLGCASGAC-----SRGRVYPMQLLEQIHKVHFLHKDT-V 437
DB 403 FVIDAYATIAHLHSMHQALCPGHTGLCPAMEPTDGRM---LLOYIRAVRENGSAGTPY 458
QY 438 AFENDRDLPLSTNIITAMD-WNGPKMTFTVLGSGTSPVOLINETKIOMHGKDNQYKSY 496
DB 459 MENENDADAGRVDIFQYQATNGSASSGGOAVGQMAET-LRLDVEALOWSGDPHEVPSSL 517
QY 497 CSSDCLGEGHORVYTGHHCCFECVPCGAGTFLNKSDLYRCOPCGKREMAPESSQCEPRT 556
DB 518 CSLPCGPRGKRKKMKVPCPCMCCEACDGYRF-QVDEFTCEACPGDMRPTPNHTGCRPTP 575
QY 557 VVFLALREHTSWVLAANTLLLLLLLG-----TAGLFAMHLDTPYVRSAGRLCEFLMG 610
DB 576 VVRL-----SWSSPWAAPRLLAVALGIVATTYVATFVRYNNTPTVRSAGRELSTVLLT 629
QY 611 SLAAGSGSLYGFGEPTRPACLRQALFALGFTIFLSCLTVRSFOLIIF--KFSKYV 667
DB 630 GIFLYATIFALWAEPAVCAARLFLGITLISALTLTKNRIYRIFEQGRSVTPP 689
QY 668 TFYHAMVVOHAGLFFMISAAOLLIC-----LFWLVVMTPLPAREYORFPHLY---- 716
DB 690 PF-----ISPTSQLVTFESLTLQVVGMIAM-LGARP-----PHSVIDYE 728
QY 717 -----MLCETENSIGFIILAFYNGLLSISAFACSYLGKDLPENYNEACVTF 764
DB 729 EORTVDEPARGVLKC-DMSDSLIGCLGYSLLIAMTYCTVYAIKARGVETFEAKPTIGF 787
QY 765 SLLENFVSWIA---FTTASYVDGKYLPAANMAGLS-SLSSGFGYFLPKCYVILCRP 819

DB 788 TMTTCTIWLAFVPIFFGTAAQSAEKIYIQTTLTVSLASASYSGLMLYVPRTYVILFHP 847
QY 820 DLN 822
DB 848 EQN 850

Search completed: June 11, 2002, 10:26:28
Job time: 471 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 09:32:26 ; Search time 43.75 Seconds
(without alignments)
1847.112 Million cell updates/sec

Title: US-09-819-946-2
Perfect score: 4493
Sequence: 1 MLCFARLVGLQLISCCWA.....NSTEHQASTIDYTRCGST 841

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1147	25.5	858	2 JC7683	taste receptor TIR
2	1142	25.4	1078	2 A56715	calcium receptor (
3	1130	25.2	1088	2 B56715	calcium receptor (
4	1126.5	25.1	1085	2 S40476	Ca(2+)-sensing rec
5	1123	25.0	1079	2 I59362	calcium/polyvalent
6	741.5	16.5	1180	2 JC2132	metabotropic gluta
7	741.5	16.5	1212	2 JC2131	metabotropic gluta
8	727	16.2	1171	2 A42916	metabotropic gluta
9	727	16.2	1199	2 A41939	metabotropic gluta
10	708.5	15.8	912	2 JH0563	G protein-coupled
11	703	15.6	1218	2 S71376	metabotropic gluta
12	692.5	15.4	872	2 JH0561	glutamate receptor
13	688	15.3	879	2 JH0562	metabotropic gluta
14	688	15.3	879	2 JH0562	metabotropic gluta
15	664	14.8	871	2 A46742	metabotropic gluta
16	652	14.5	908	2 A49142	metabotropic gluta
17	619.5	13.8	915	2 A49874	metabotropic gluta
18	590.5	13.1	999	2 T27628	hypothetical prote
19	517.5	11.5	1267	2 T21340	hypothetical prote
20	398	8.9	551	2 T30806	metabotropic gluta
21	263	5.9	1099	2 T16283	metabotropic gluta
22	254	5.7	923	2 F84732	hypothetical prote
23	253	5.6	921	2 T51136	probable ligand-ga
24	218	4.9	1039	2 T45779	ionotropic gluta
25	192.5	4.3	933	2 C96495	probable glutamate
26	192	4.3	925	2 T06128	probable ligand-ga
27	182	4.1	953	2 E84732	hypothetical prote
28	181.5	4.0	958	2 T02741	probable ligand-ga
29	178.5	4.0	925	2 T51133	ligand gated chann

30	178.5	4.0	951	2 T51132	probable glutamate
31	178.5	4.0	975	2 A84550	probable ligand-ga
32	178	4.0	912	2 T51131	ligand gated chann
33	178	4.0	997	2 S33754	glutamate receptor
34	176	3.9	960	2 JE0356	gamma-aminobutylic
35	169.5	3.8	950	2 T51134	ionotropic gluta
36	166	3.7	962	2 D86186	hypothetical prote
37	165	3.7	965	2 T51244	N-methyl-D-asparta
38	163	3.6	938	2 A46612	N-methyl-D-asparta
39	163	3.6	976	2 T51137	ionotropic gluta
40	160	3.6	941	2 T51135	ligand-gated chann
41	158	3.5	885	2 JN0339	N-methyl-D-asparta
42	158	3.5	901	2 JN0337	N-methyl-D-asparta
43	158	3.5	922	2 JN0338	N-methyl-D-asparta
44	158	3.5	938	2 S19710	N-methyl-D-asparta
45	156	3.5	938	2 S21104	N-methyl-D-asparta

ALIGNMENTS

RESULT	1	Query Match	25.5%	Score 1147;	DB 2;	Length 858;
JC7683	1	Best local similarity 32.6%; Pred. No. 2e-79;				
Sequence: Mus musculus (house mouse)		Matches 271; Conservative 150; Mismatches 356; Indels 54; Gaps 16;				
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001						
C/Accession: JC7683						
R:Kitagawa, M.; Kusakabe, Y.; Miura, H.; Ninomiya, Y.; Hino, A.						
Biochem. Biophys. Res. Commun. 283, 236-242, 2001						
A>Title: Molecular genetic identification of a candidate receptor gene for sweet tast						
A/Reference number: JC7683; MUID:21222875; PMID:11322794						
A/Contents: tongue						
A/Accession: JC7683						
A/Molecule type: mRNA						
A/Residues: 1-858 <KIT>						
A/Cross-references: DBJ:AB049994						
C/Comment: This protein, a seven-transmembrane receptor, serves as a mediator of the						
A/Genes: tlr3						
A/Map position: 4						
C/Keywords: transmembrane protein						
Query Match						
Best local similarity 32.6%; Pred. No. 2e-79;						
Matches 271; Conservative 150; Mismatches 356; Indels 54; Gaps 16;						
QY 29 SPDFTLPGDYLLAGLFPPLHS--GCLQVRHREPTLDCRSCSFNKGHYLFOAMRLGVEE 85						
DB 26 SQQFAAGDYLLGFLPGSTREATLNRQTPNSIPCNR---FSPLGFLFAMAKMAVEE 82						
QY 86 INNSTALLPNTLLGQYLDVCSDS-ANYATLRVLSLGGQHIEIQLGDLKSPVYAVI 144						
DB 83 INNGSALLPGLRGLGDLDTCEPVTYMKSSIMFLAKYSOSIAYCVYTOQOPRVAVI 142						
QY 145 GPDSNRAATTAALSPVLPVPIVSAASSETLSVROYSPFLRTIPNKKYQVETWLLQ 204						
DB 143 GPHSEALLITGKFFSFLLMPQVSTASASMDRLSDRETPSPFRITYPSDRVQLQAVVTLQ 202						
QY 205 KFGWTWISLVGSSDDYGLQVQVALENQATGQICIAFDIMEFSAQVDE--RMQCLMRH 262						
DB 203 NFSNMVVAALGSDDDYGGREGLSIFSSLANARQICIAHGIVQHDTSQQLGKLVDLRQ 262						
QY 263 LAQAGATVYVYFSSQQLARVFESVVLNLNKGKVVAVASANAALSHITGVQIQTIGVLT 322						
DB 263 VNQSIVQVYVYFASARAVYSLEYSYIHGGLSPKVVAVASESWLTSVLTLPINAVGVYL 322						
QY 323 GVAIQKRAVPGKAFKAEVYARADKAPRCHKSGWSS-----NOLCREQAF 370						
DB 323 GFLQKALIPERSHYVETHAL---AADP-----AFCAQLMELDEEHVNGRCRCDDI 375						
QY 371 MAHTPKLKAFSSMSA-----YNNAYRAVYAVAHGLHQLLGASGAGCS--RGVYVPMOLLE 423						

Db 376 MTQNSGILLQNLISAGLQHQTFAITYAAYVSAQAALHNTLQCNVSHCHVSEHYLPQMLE 435
 Qy 424 QIHKVYFLHKPTVAENDNRDPLSSYNIITAMPNGPKMTFTVYSGSTWSPVOLINETKI 483
 Db 436 NMYNMSFHAARDLLOFDAGNDVMEYDLKMWQSPPTVLHYGTNGT---LQLOQSKM 492
 Qy 484 QNHKDNVPKSVCSDDLGHQRVVTFHHCCEFCVPGAGTFLNKSDLYRCQPCGKEE 543
 Db 493 YWPG--NQVPVSCSRQCKDQVRRKGFHSCCYDCVDCAGSYRKHPDDFTCTPCNQDQ 550
 Qy 544 WAPESQCFEPRTVFLAREHTSWLLAANTLLLLLTGTLGFMHIDTPVRSAGG- 602
 Db 551 WSPENSTACLPKPKFLANGEVPLVSLLLLCIVLGLAALGLSVHMDSPVQAQSGS 610
 Qy 603 RLCE--LIMGSLAAGSGSYGFEGETPRACLLRQALFALGFTFLSCLTFRSFLITIF 660
 Db 611 QRCFGLICGLFCL---SVLLFPGRRSSASCLAQPMALPLTGCLSTFLQAAEFV-- 665
 Qy 661 KESTKVPRTYHAMVONHAGLF---VMISAAQLLICLTWLVWTPLPAREYQRPBHV 716
 Db 666 --ESELPLSMAMWLCISYLRGLMAMLVLLATFEVLAALCAMYLIAFPEVYVDMVSLPTEV 723
 Qy 717 MECEFTNSLGLTFLAFLYGLISAFASYLGKDPENYNEAKCTESLLEFVSWIAF 776
 Db 724 LECHVRSWVSLGLVHITAMLAFLCIGTFLVQSQPGRYNRARGLTFMLAYFTIWSF 783
 Qy 777 FTTASVYDCKYLPAAANMAGLSLSSGFGYFLPKCYVLCRPDLNTEHEF 827
 Db 784 VPLLANVQVAYQPAVOMGALVLCALGILVTFHLPKCYVLLMPLKMTQGEF 834
 RESULT 2
 A56715
 calcium receptor (clone pHPCar-4.0) - human
 C.Species: Homo sapiens (man)
 C.Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 01-Dec-2000
 C.Accession: A56715; S49341; A49419; B49419; C49419
 R.Garrett, J.E.; Capuano, I.V.; Hammerland, L.G.; Hung, B.C.P.; Brown, E.M.; Hebert, S.C.
 J. Biol. Chem. 270, 12919-12925, 1995
 A.Title: Molecular cloning and functional expression of human parathyroid calcium recept
 A.Reference number: A56715; MUID:95279439
 A.Accession: A56715
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-1078 <GAR>
 A.Cross-references: GB:U20759; NID:9683744; PIDN:AAA86503.1; PID:9683745
 R.Pearce, S.H.S.; Thakker, R.V.
 submitted to the EMBL Data Library, August 1994
 A.Reference number: S49341
 A.Accession: S49341
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-180, 'Q', 182-989, 'R', 991-1078 <PEA>
 A.Cross-references: EMBL:X81086
 R.Pollak, M.R.; Brown, E.M.; Chou, Y.H.; Hebert, S.C.; Marx, S.J.; Steinmann, B.; Levi,
 Cell 75, 1297-1303, 1993
 A.Title: Mutations in the human Ca(2+)-sensing receptor gene cause familial hypocalciuric
 A.Reference number: A49419; MUID:94094324
 A.Accession: A49419
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 178-192 <POL>
 A.Experimental source: family N
 A.Note: sequence inconsistent with nucleotide translation
 A.Note: sequence modified after extraction from NCBI backbone
 A.Note: 186-Arg mutation is associated with familial hypocalciuric hypercalcemia and nec
 A.Note: sequence extracted from NCBI backbone (NCBIN:142453)
 A.Accession: B49419
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 289-303 <PO2>
 A.Experimental source: family E
 A.Note: sequence modified after extraction from NCBI backbone

A.Note: 298-Lys mutation is associated with familial hypocalciuric hypercalcemia and
 A.Note: sequence extracted from NCBI backbone (NCBIN:142455)
 A.Accession: C49419
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 788-802 <PO3>
 A.Experimental source: family J
 A.Note: sequence modified after extraction from NCBI backbone
 A.Note: 796-Trip mutation is associated with familial hypocalciuric hypercalcemia and
 A.Note: sequence extracted from NCBI backbone (NCBIN:142457)
 C.Keywords: glycoprotein; receptor; transmembrane protein

Query Match 25.4%; Score 1142; DB 2; Length 1078;
 Best local similarity 31.0%; Pred. No. 6.3e-79;
 Matches 276; Conservative 163; Mismatches 358; Indels 92; Gaps 22;

Qy 16 SCCW--AFACHSTESSPD--FTLPEDYLLAGFLPHSLCL---QVRRPEVTLCDRSC 66
 Db 5 SCCWVLLATLWHTSAVGPQRAQKKGDIILGGLFPIHFGVAARDKDDLSRPSVEICR-- 62
 Qy 67 SENEHGYHLFOAMRUGVEEINNSTALLPNTLGYOLDYDCSDSANVYATLFLVSLPGOH 126
 Db 63 -TNEGFRLQAMIFALIEINSSPALPNTLGYRLEFCNTVSK--ALEATLSEYAQNK 119
 Qy 127 IELOG-----DLIHSPTVLAVIPDSTINRAATTAALSPFLVPMISYASSETLSVKRQ 181
 Db 120 IDSINDEFRCNSEHPTSLAVAGATSGSVTAVALLLFLFIPQVYASASSRLLSNKKQ 179
 Qy 182 YPSFLRTINDKXQVETMYLLLOKRGWYISLVGSSDDYGOGLVQALENOATGOGICLAF 241
 Db 180 FKSFLRTINDKQVAMADIIEFRMNVGTIAADDYGRGIEKFRREAEERDIDCF 239
 Qy 242 KDIIMPQAQVDEBMOCLRHIAQAGATVYVYSSRQLARVEESVLTNLGKVVASE 301
 Db 240 SEIL--SQSDEBEIOHVEVTONSTAKIYVFPSSPDLEPLKEIVRNITGKIMLASE 297
 Qy 302 AMALSHRITGVPGIORIGVLAIOKRAVPGIKAF-----EEAYARADKEAPRP 351
 Db 298 AMASSSLIAMPOVFHVGGTIGFALKAGQIPGFRFLKRVHPRKSVHNGFAKEFMEETPN 357
 Qy 352 CH-----KG-----SWCSSNOLCEBQAFAMHPKAKAFMSMSA 386
 Db 358 CHLOEGAKPLPYDFFLRGHESGDRFSNSTAFRLCTGDEISSVEPTDYTHLRIS 417
 Qy 387 YNAAYRAVAVAGHLOLLC-----ASGACSR--GRVPMOLEQIHKVHFLHK--DTV 437
 Db 418 YNVIYLAIVSIAHRLDITYCLGRGLFTNGSCADIKKVAQVILKRLRLNLTNNNGEDV 477
 Qy 438 AFNDNRDPLSSYNIITAMPNGPKMTFTVYSGSTWSPV-----OLINETKIOMH 486
 Db 478 TFECDGLGVNGYSIIMHLSPED-----GSIVFKVGYNYAKKGERLFTNEEKILMS 531
 Qy 487 GKDQVPRKSVCSDDLCEGHQR--VYGFHHCCEFCVPGAGTFLNKSDLYRCQPCGKEEA 545
 Db 532 GSRREVPFNCGRCDLAGTRKGIIEGEPCCFCEVCPGEVSDASACNKPDPDWS 591
 Qy 546 PEGSOTCFPRTVFLARHTSWVLLAANTLLLLLTGTLGFMHIDTPVRSAGGRIC 605
 Db 592 NENHTSIAKEIEFLSWTEPFGALTLFLFAVLGFLAFYIKRNPPIYKATRELS 651
 Qy 606 FLMLGSLAAGSGSYGFEGETPRACLLRQALFALGFTFLSCLTFRSFLITIFKFSK 665
 Db 652 YLLFELLCFSSSLEFIEPDQMCRLQPAFGISFVICIILVKTIRVLVEE--AK 709
 Qy 666 VPTFYH--AMVONHAGLFMMISSAQLLICLTWLVWTPLPAREYQRPBHVLMLEETEN 724
 Db 710 IPTSFHRKWWGILNLOFLVFLCTFMQIVCVILVLTAPSPSRNOLDEDELIITFHES 769
 Qy 725 --SLGFIILAVNGLSISAFASYLGKDPENYNEAKCTESLLEFVSWIAFTT--AS 781
 Db 770 LMAIGFLIG--YTCLLAALCFEFAFRKRLPENFNNAKITTSMLIFFIWSIFPAVMS 827

QY 782 VYDKRYLPANMAGLSLSSGFGYFLPKCYVILCRPDLNSTHFQAS 830
 B56715
 Db 828 TY-GKFSVAVEVIAIILASFGLLACIFPNKIYILLFKPSRMTIEVRCS 875

RESULT 3

calcium receptor (clone phPCAR-5.2) - human

C/Species: Homo sapiens (man)
 C/Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000

C/Accession: B56715
 R/Garret, J.E.; Capriano, I.V.; Hammerland, L.G.; Hung, B.C.P.; Brown, E.M.; Hebert, S.C.
 J. Biol. Chem. 270, 12919-12925, 1995
 A/Title: Molecular cloning and functional expression of human parathyroid calcium receptor
 A/Reference number: A56715; MUID:95279439
 A/Accession: B56715
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-1088 <GAR>
 A/Cross-references: GB:U20760; NID:9683746; PIDN:AAA86504.1; PID:9683747

C/Keywords: glycoprotein; receptor; transmembrane protein

Query Match 25.2%; Score 1130; DB 2; Length 1088;
 Best Local Similarity 30.8%; Pred. No. 5.3e-78;
 Matches 277; Conservative 162; Mismatches 358; Indels 102; Gaps 23;

QY 16 SCCM---AFACHSTESSPD--FTLPDYLALGLPPLHSGCL---QVRHREVTLCRSC 66
 B56715
 Db 5 SCCVLLALMTWTSAYGPDORAKKGDIIIGLFPPIHGVAVAKDOKDLSRESVECTR-- 62
 QY 67 SFNEGHYLFQAMRLGVEEINNSTALLPNITLGYOLYVCSDSANVYATLRLSPGOH 126
 Db 63 -YNRGRWLOAMFAIEEINSSPALLPNLTLGRITDCTVSK--ALEATLSFVAONK 119
 QY 127 IELOG----DLHSPVTLAVIGPDSNTRATTAALISPLVPMISYASSETLSYVKRQ 181
 Db 120 IDSLNLFDFNCSEHISTIAVAVATGSGVSTAVANLGLFYIPQVSAASSRLSNKQ 179
 QY 182 YPSFLRTIPNDKYQVETWVLLQKFGMTWISLVGSSDDYQGLGVQALENQTGGICIAF 241
 Db 180 FKSLRTIPNDHQAADIIIEFRMNVGTIAADDDYGRPGIEKREKEEEDIDICDF 239
 QY 242 KDIPFAQVDERMOCLMRLHAOGATVVVVFSSROLARVFESVLTNLTKGVWASE 301
 Db 240 SELT--SOYSDERIOHVEYIONSTAKVIYVFSSGPDLEPLKEIYARNITGKIWLASE 297
 QY 302 AMALSRHTGVPGIQRIGMVGVAIQRKAVGLKAF-----EEAVARADKEARP 351
 Db 298 AMASSSLIAMPQYPHVVGTFGFAKAGQIPGFRFLKTVHPRKSVHNGFAKEEETFN 357
 QY 352 CH-----KG-----SWCSSNOLCREQAMATMPKLFKAFSSSA 386
 Db 358 CHLOEAGKPLPVDFTLRGHEESGDRFSNSTARFPLCTGGENISSVETPIIDYTHLIS 417
 QY 387 YNAVAYAVAHGLHQLG-----ASGACSR--GRVYPMQLLEQIHKVFLLHK--DTV 437
 Db 418 YNVTLAYISIAHALODIYTCIPGRGLFTNGSCADIKKVEAMOVKHLHLNFTNMGOV 477
 QY 438 AFNDNRPLSSYNTIADWNGPKMTFTVLGSSVSPV-----QLNINEFKIQWH 486
 Db 478 TFDECGDLVGYSTIINMHLSPED-----GSIVFEKVGYYNVYAKKGRFLINNEKILMS 531
 QY 487 GKDN-----QVRKSVSSDCLBQHQR--VVTGFHHCCEVPCGAGTFLKSLDYR 535
 Db 532 GFSEPLFLVSVLQVFPNSCRSLAGTRKGIIEGPEPTCFCEVCEPDGEYSDETASA 591
 QY 536 CQPCGKEMAPESQTCFPRVYVFLALREHNSWLLAANTLLLLGLTAGLFLMHIDTP 595
 Db 592 CNKCPDDPMSNENTSCIAKEIFLSWTERPGIALFLPAVGIPLAFVLGAVTFKFRPTP 651
 QY 596 VRSAGRLCFMLGSLAAGSGSLYGFGEPTRPACLLRQALFALGTFILFSLCTLVASFQ 655

Db 652 IVKATNRELSTLLFSLCCFSSSLFFIGEPQDWTCLRQPAFGISFVLCISCLIVKTR 711
 QY 656 LIIFKSTKPTPEYH--AWQNHGAGLFVMTSSAOLICLTWLVMPPLAREQORPH 714
 Db 712 VLVVE--AKIPSTFHRKMGWLNLOFLVFLCTMQIYICVIMLYTAPSSYRQOELEDE 769
 QY 715 LVMLECTETN--SLGFIILAFYNGLLSISAFACSYLQKDLPENYNEAKVTFSLFNEVS 772
 Db 770 IIFITCHEGSLMAGFLIG--YTCLLAIACFFPAFKSKRLPENNEAKFIIFMSILFIYI 827
 QY 773 WIAFPTT--ASYDDKYLPANMAGLSLSSGFGYFLPKCYVILCRPDLNSTHFQAS 830
 Db 828 WISFIPAVASTY--GKFSVAVEVIAIILASFGLLACIFPNKIYILLFKPSRMTIEVRCS 885

RESULT 4

S40476
 Ca(2+)-sensing receptor - bovine

C/Species: Bos primigenius taurus (cattle)
 C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999

C/Accession: S40476
 R/Brown, E.M.; Gamba, G.; Riccardi, D.; Lombardi, M.; Butters, R.; Kilfor, O.; Sun, A.
 Nature 366, 575-580, 1993
 A/Title: Cloning and characterization of an extracellular Ca(2+)-sensing receptor fro
 A/Reference number: S40476; MUID:94077182

A/Accession: S40476
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-1085 <BRO>
 A/Cross-references: GB:S67307; NID:9453108; PIDN:AAB29171.1; PID:9453109

Query Match 25.1%; Score 1126.5; DB 2; Length 1085;
 Best Local Similarity 30.7%; Pred. No. 9.8e-78;
 Matches 274; Conservative 163; Mismatches 362; Indels 93; Gaps 22;

QY 14 LISCCMAFACHST---ESSPD--FTLPDYLALGLPPLHSGCL---QVRHREVTLCRSC 63
 Db 3 LYSCWILAFSLTMTCTSAVGPDOAKKGDIIIGLFPPIHGVAVAKDOKDLSRESVECTR 62
 QY 64 RSCSFNEGHYLFQAMRLGVEEINNSTALLPNITLGYOLYVCSDSANVYATLRLSPGOH 123
 Db 63 R--YNRGRWLOAMFAIEEINSSPALLPNLTLGRITDCTVSK--ALEATLSFVA 117
 QY 124 OHTELOG----DLHSPVTLAVIGPDSNTRATTAALISPLVPMISYASSETLSY 178
 Db 118 ONKIDSLNLFDFNCSEHISTIAVAVATGSGISTAVANLGLFYIPQVSAASSRLSN 177
 QY 179 KROYPSFLRTIPNDKYQVETWVLLQKFGMTWISLVGSSDDYQGLGVQALENQTGGIC 238
 Db 178 KNQFKSFLRTIPNDHQAADIIIEFRMNVGTIAADDDYGRPGIEKREKEEEDIDIC 237
 QY 239 IAFKDIMEFSAQVDERMOCLMRLHAOGATVVVVFSSROLARVFESVLTNLTKGVW 298
 Db 238 IDFSLEI--SOYSDERIOHVEYIONSTAKVIYVFSSGPDLEPLKEIYARNITGRIWL 295
 QY 299 ASEAMALSRHTGVPGIQRIGMVGVAIQRKAVGLKAF-----EEAVARADKEA 348
 Db 296 ASEAMASSSLIAMPDEYFHVVGTFGFAKAGQIPGFRFLKTVHPRKSVHNGFAKEE 355
 QY 349 PRPCH-----KGSWSSNOLCREQAMATMPKLFKAFSSSA 383
 Db 356 TFNCHLOEAGKPLPVDFTLRGHEEGGKARLSNPFARFPLCTGGENISSVETPIIDYTHL 415
 QY 384 SSAYNAVAYAVAHGLHQLG-----ASGACSR--GRVYPMQLLEQIHKVFLLHK-- 434
 Db 416 RISTNVYLAISIAHALODIYTCIPGRGLFTNGSCADIKKVEAMOVKHLHLNFTNM 475
 QY 435 DTVAFNDNRPLSSYNTIADWNGPKMTFTVLGSSVSPV-----QLNINEFKI 483
 Db 476 EOVTFDECGDLVGYSTIINMHLSPED-----GSIVFEKVGYYNVYAKKGRFLINDEKI 529
 QY 484 QWHGKDNQVPRKSVSSDCLBQHQR--VVTGFHHCCEVPCGAGTFLKSLDYRQOPGCKE 542

[illegible]

RESULT 5

159362

calcium/polyvalent cation-sensing receptor precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Jul-1996 #sequence-Revision 02-Jul-1996 #text-change 05-Nov-1999

C:Accession: 159362; A55594

R:Rut, M.; Molliver, M.E.; Snowman, A.M.; Snyder, S.H.

Proc. Natl. Acad. Sci. U.S.A. 92, 3161-3165, 1995

A:Title: Calcium sensing receptor: molecular cloning in rat and localization to nerve terminals

A:Reference number: 159362; MUID:95241465

A:Accession: 159362

F:187-212/Region: hydrophobic

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1079 <RES>

A:Cross-references: EMBL:U02089; NID:G790578; PIDN:AA52195.1; PID:G790579

A:Experimental source: striatal

R:Riccardi, D.; Park, J.; Lee, W.S.; Gamba, G.; Brown, E.M.; Hebert, S.C.

Proc. Natl. Acad. Sci. U.S.A. 92, 131-135, 1995

A:Title: Cloning and functional expression of a rat kidney extracellular calcium/polyvalent cation-sensing receptor

A:Reference number: A55594; MUID:95116508

A:Accession: A55594

A:Molecule type: mRNA

A:Residues: 1-133, 'X', 135-1079 <RIC>

A:Cross-references: GB:U10354

A:Experimental source: kidney

A:Keywords: calcium; glycoprotein; phosphoprotein; transmembrane protein

F:1-20/Domain: signal sequence #status predicted <SIG>

F:187-212/Region: hydrophobic

F:613-635/Domain: transmembrane #status predicted <TM1>

F:650-670/Domain: transmembrane #status predicted <TM2>

F:683-700/Domain: transmembrane #status predicted <TM3>

F:725-744/Domain: transmembrane #status predicted <TM4>

F:770-790/Domain: transmembrane #status predicted <TM5>

F:806-828/Domain: transmembrane #status predicted <TM6>

F:841-860/Domain: transmembrane #status predicted <TM7>

F:900-921/Region: hydrophobic

F:921-940/Domain: transmembrane #status predicted <TM8>

F:940-960/Domain: transmembrane #status predicted <TM9>

F:960-980/Domain: transmembrane #status predicted <TM10>

F:980-1000/Domain: transmembrane #status predicted <TM11>

F:1000-1020/Domain: transmembrane #status predicted <TM12>

F:1020-1040/Domain: transmembrane #status predicted <TM13>

F:1040-1060/Domain: transmembrane #status predicted <TM14>

F:1060-1080/Domain: transmembrane #status predicted <TM15>

F:1080-1100/Domain: transmembrane #status predicted <TM16>

F:1100-1120/Domain: transmembrane #status predicted <TM17>

F:1120-1140/Domain: transmembrane #status predicted <TM18>

F:1140-1160/Domain: transmembrane #status predicted <TM19>

F:1160-1180/Domain: transmembrane #status predicted <TM20>

F:1180-1200/Domain: transmembrane #status predicted <TM21>

F:1200-1220/Domain: transmembrane #status predicted <TM22>

F:1220-1240/Domain: transmembrane #status predicted <TM23>

F:1240-1260/Domain: transmembrane #status predicted <TM24>

F:1260-1280/Domain: transmembrane #status predicted <TM25>

F:1280-1300/Domain: transmembrane #status predicted <TM26>

F:1300-1320/Domain: transmembrane #status predicted <TM27>

F:1320-1340/Domain: transmembrane #status predicted <TM28>

F:1340-1360/Domain: transmembrane #status predicted <TM29>

F:1360-1380/Domain: transmembrane #status predicted <TM30>

F:1380-1400/Domain: transmembrane #status predicted <TM31>

F:1400-1420/Domain: transmembrane #status predicted <TM32>

F:1420-1440/Domain: transmembrane #status predicted <TM33>

F:1440-1460/Domain: transmembrane #status predicted <TM34>

F:1460-1480/Domain: transmembrane #status predicted <TM35>

F:1480-1500/Domain: transmembrane #status predicted <TM36>

F:1500-1520/Domain: transmembrane #status predicted <TM37>

F:1520-1540/Domain: transmembrane #status predicted <TM38>

F:1540-1560/Domain: transmembrane #status predicted <TM39>

F:1560-1580/Domain: transmembrane #status predicted <TM40>

F:1580-1600/Domain: transmembrane #status predicted <TM41>

F:1600-1620/Domain: transmembrane #status predicted <TM42>

F:1620-1640/Domain: transmembrane #status predicted <TM43>

F:1640-1660/Domain: transmembrane #status predicted <TM44>

F:1660-1680/Domain: transmembrane #status predicted <TM45>

F:1680-1700/Domain: transmembrane #status predicted <TM46>

F:1700-1720/Domain: transmembrane #status predicted <TM47>

F:1720-1740/Domain: transmembrane #status predicted <TM48>

F:1740-1760/Domain: transmembrane #status predicted <TM49>

F:1760-1780/Domain: transmembrane #status predicted <TM50>

F:1780-1800/Domain: transmembrane #status predicted <TM51>

F:1800-1820/Domain: transmembrane #status predicted <TM52>

F:1820-1840/Domain: transmembrane #status predicted <TM53>

F:1840-1860/Domain: transmembrane #status predicted <TM54>

F:1860-1880/Domain: transmembrane #status predicted <TM55>

F:1880-1900/Domain: transmembrane #status predicted <TM56>

F:1900-1920/Domain: transmembrane #status predicted <TM57>

F:1920-1940/Domain: transmembrane #status predicted <TM58>

F:1940-1960/Domain: transmembrane #status predicted <TM59>

F:1960-1980/Domain: transmembrane #status predicted <TM60>

F:1980-2000/Domain: transmembrane #status predicted <TM61>

F:2000-2020/Domain: transmembrane #status predicted <TM62>

F:2020-2040/Domain: transmembrane #status predicted <TM63>

F:2040-2060/Domain: transmembrane #status predicted <TM64>

F:2060-2080/Domain: transmembrane #status predicted <TM65>

F:2080-2100/Domain: transmembrane #status predicted <TM66>

F:2100-2120/Domain: transmembrane #status predicted <TM67>

F:2120-2140/Domain: transmembrane #status predicted <TM68>

F:2140-2160/Domain: transmembrane #status predicted <TM69>

F:2160-2180/Domain: transmembrane #status predicted <TM70>

F:2180-2200/Domain: transmembrane #status predicted <TM71>

F:2200-2220/Domain: transmembrane #status predicted <TM72>

F:2220-2240/Domain: transmembrane #status predicted <TM73>

F:2240-2260/Domain: transmembrane #status predicted <TM74>

F:2260-2280/Domain: transmembrane #status predicted <TM75>

F:2280-2300/Domain: transmembrane #status predicted <TM76>

F:2300-2320/Domain: transmembrane #status predicted <TM77>

F:2320-2340/Domain: transmembrane #status predicted <TM78>

F:2340-2360/Domain: transmembrane #status predicted <TM79>

F:2360-2380/Domain: transmembrane #status predicted <TM80>

F:2380-2400/Domain: transmembrane #status predicted <TM81>

F:2400-2420/Domain: transmembrane #status predicted <TM82>

F:2420-2440/Domain: transmembrane #status predicted <TM83>

F:2440-2460/Domain: transmembrane #status predicted <TM84>

F:2460-2480/Domain: transmembrane #status predicted <TM85>

F:2480-2500/Domain: transmembrane #status predicted <TM86>

F:2500-2520/Domain: transmembrane #status predicted <TM87>

F:2520-2540/Domain: transmembrane #status predicted <TM88>

F:2540-2560/Domain: transmembrane #status predicted <TM89>

F:2560-2580/Domain: transmembrane #status predicted <TM90>

F:2580-2600/Domain: transmembrane #status predicted <TM91>

F:2600-2620/Domain: transmembrane #status predicted <TM92>

F:2620-2640/Domain: transmembrane #status predicted <TM93>

F:2640-2660/Domain: transmembrane #status predicted <TM94>

F:2660-2680/Domain: transmembrane #status predicted <TM95>

F:2680-2700/Domain: transmembrane #status predicted <TM96>

F:2700-2720/Domain: transmembrane #status predicted <TM97>

F:2720-2740/Domain: transmembrane #status predicted <TM98>

F:2740-2760/Domain: transmembrane #status predicted <TM99>

F:2760-2780/Domain: transmembrane #status predicted <TM100>

F:2780-2800/Domain: transmembrane #status predicted <TM101>

F:2800-2820/Domain: transmembrane #status predicted <TM102>

F:2820-2840/Domain: transmembrane #status predicted <TM103>

F:2840-2860/Domain: transmembrane #status predicted <TM104>

F:2860-2880/Domain: transmembrane #status predicted <TM105>

F:2880-2900/Domain: transmembrane #status predicted <TM106>

F:2900-2920/Domain: transmembrane #status predicted <TM107>

F:2920-2940/Domain: transmembrane #status predicted <TM108>

F:2940-2960/Domain: transmembrane #status predicted <TM109>

F:2960-2980/Domain: transmembrane #status predicted <TM110>

F:2980-3000/Domain: transmembrane #status predicted <

Query Match	25.0%;	Score 1123;	DB 2;	Length 1079;
Best Local Similarity	30.9%;	Pred. No. 1.8e-77;		
Matches 275;	Conservative 162;	Mismatches 360;	Indels 92;	Gaps 22

```

0y 16 SCC---WAFACHSTESSPD--FTLPBGDYLAGLPPLHSGCL---QVRHREYTLCDRSC 66
    |||  |  |||  |  |||  |  |||  |  |||  |  |||  |  |||  |  |||  |
Db 5  SCCLALLALAHWSSAYGPDQRAQKGGDIILGCLPPIHFGVAAKQDJDKSRSEYECIR-- 62

0y 67 SPNEHGCHLPQAMRLGVEEINNSTALLPNTTLGQLYQOLYDVCSDSANVYATATLRVLSLPGOH 126

```

Db	63	-YNRGGRWGLQAMIFALEELINSSPLLNPNTLQRIFFDCNTRYK--ALEATLSPVAQNK	119
Qy	127	IELQG-----DLHYSPTVLAVIGPDSTNRAATTAALSPLFLPMISVAASETJLSVKRQ	181
Db	120	IDSJNLDEFCNCSBHIKSTJIAVVGATGSGSVTAVMILGIFYIPQVSYASSRLLSNKQ	179
Qy	182	YPSFLRTIPNDKYVEYVEMVLLQKFGWTYJLSLVGSSDDYGGOLGYQALNEMQATGOGCIATF	241
Db	180	YKSLRTIPIPDHEQATMAOIEEFYFNNWVGTIAADDDGPRGIEKFEFREAEERICIDFE	239
Qy	242	KDIMEFSAQVDEEMQCLMHLMAQAGTAVVVPSSROLKRVFEESVJLNLGKWAVASE	301
Db	240	SELI--SQYSDEETIQOYVEYIÖNSTAKTYIVYSSGPDLBPLKEIVKRNITGRITWLSE	297
Qy	302	AMALSRIITGVPGIÖRIGWLVGAALOKRAVPYGLKAF-----EEAYARADKEARP	351
Db	298	AMASSSLIAPEYHYVVGITIGFKLAKGQJLPGREFLQKVHPRKSVHNOFAKEFMEDEFEN	357
Qy	352	CH-----KG-----SWCSSNOLCCEQCAFMAHTPKATKASMSA	386
Db	358	CHLOEGAKGLPYDTEVFRSHSEGGNRLNLSSTFRPLCTGDEINISVENPYMDYELHRTS	417
Qy	387	YNARAYAAVAHGHJOLG-----ASAGACR--GRVYPMOLIEÖHKVHFLHK--DHY	437
Db	418	YNYVLAYASTAHALQDIYTCLPGRGLEFLTSSCADIKKVEAMÖVLKHLRLNFTNNNGEYV	477
Qy	438	AFNNRPDLPLSSYNIADMDMGPKWTFEVLSSYTSWSPV-----OLANETKIOWH	486
Db	478	TFDECGDLGVNGYSIINHLPED-----GSIFYKEVGYNYAKKGEHLPIINEKLIMS	531
Qy	487	GKDNQVPKSVCCSDCLEGHOR--VYTGHHCCPECPVPCGAGTFLNKSDLYRCOPCGKEEMA	545
Db	532	GFSREYFVSNCRQCAQTKRKGJIEGEPITCCFCEVCPCPGESGGETDASACDKCDDPWS	591
Qy	546	PEGQOTEPFRVYVLAJREHTSWLLAANTLLLLLLGTAJGFAMHLDOPVYRSAGRJC	605
Db	592	NEHNTSCIAKITEPLATEPBGJALTLFAVLGJFLTAFLVGVAFIKRRNPDIYKATNRELS	651
Qy	606	FLMGLSLAAGSGSLYGFGEPTRPACLRQALPALGFTTFLGSLIVRSQJLITIKFSTK	665
Db	652	YLLFLSLCCSSSLFETIGERQDMTCRLRPAGISFVLCISILVKTRVLLVE--AK	709
Qy	666	VPTEYH--AMVONHGAGLFVMISSAÖLLICLWLVVYTPLPAREYQRFHVLMELECTETN	724
Db	710	IPTSFHRKMGMLNÖFLVLELCTFMÖLLICIDWLYAPBSYNNHLEDEJEIFTICHEGS	769
Qy	725	--SLGFLIAFLYNÖLJETSAPACSJYGLKDLPEYNNAKCVTBSLNFENPSYATFFTY--AS	781
Db	770	IMAQGSILG--YTÖLLAAICFFFAFKSRKJLPENFNNAKFTJSMLLFFIYWIJSIFIPAYAS	827
Qy	782	YVDKGYLPAANMMAGLSSLSGFGYVLPYPCYVILÖRPDLNSENHQAS	830
Db	828	TY--GRFVSAAVEVIAILLASGGLACIFENKVYIILKPSRPNITIEVYRSS	875

```

RESULT      6
JC2132
metabotropic glutamate receptor 5 A - human
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 05-Jan-1996
C:Accession: JC2132
R:Matkamlı, R.; Katsuki, F.; Yamamoto, T.; Nakamura, K.; Sugiyama, H.
Biochem. Biophys. Res. Commun. 199, 1136-1143, 1994
A:Title: Molecular cloning and the functional expression of two isoforms of
A:Reference number: JC2131; WUID:94197656
A:Accession: JC2132
A:Molecule type: mRNA

```

C:Comment: This protein is coupled to guanine nucleotide binding proteins.
C:Keywords: glycoprotein; neurotransmitter; receptor; transmembrane protein
F:560-604/Domain: transmembrane #status predicted <TM>
F:617-637/Domain: transmembrane #status predicted <TM2>

Db 647 ORIGIGLSPAMSYALVTKTNRRIARILAGSKKICTKRPMNSACAQIVIAFLICI--- 703

Qy 689 AOLLCITLWVMTPLPAREYORPHL--VMECTETNSIGFLIAFLYNGLSISAFACS 746

Db 704 -QLGIIVAFIMERPDIIMDY---PSIREVYLICNTTN-LGVYTPGLYNGLLISTCFYFA 758

Qy 747 YIGKDLPENYNEAKCVTFSLFNFSWIAFFTASYDGKYLPAANMAGLS-SLSSGFG 805

Db 759 EKTRNVPANFNENAKYIAFTVYTCIIMLAF---VPYFGSNYKIIIMCFSVSLSATVALG 815

Qy 806 GYELPKCYVILCRPDNSTEHFQAS 830

Db 816 CMFPRVYIILAKPERNWSAFTTS 840

RESULT 8

A42916 metabotropic glutamate receptor mglur5 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: A42916

R:Abbe, T.; Sugihara, H.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakanishi, S.

J. Biol. Chem. 267, 13361-13368, 1992

A:Title: Molecular characterization of a novel metabotropic glutamate receptor mglur5

A:Reference number: A42916; MUID:92317054

A:Accession: A42916

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-1171 <ABE>

A:Cross-references: GB:DI0891; NID:9220813; PIDN:BA01711.1; PID:d1002186; PID:g220814

A:Experimental source: brain

A>Note: Sequence extracted from NCBI backbone (NCBIN:107749, NCBIPI:107750)

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.2%; Score 727; DB 2; Length 1171;

Best Local Similarity 27.2%; Pred. No. 3.7e-47;

Matches 235; Conservative 151; Mismatches 356; Indels 122; Gaps 33;

Qy 34 LPGDYLLAGLFPPLHSGCLQVRRHREY-TLCDRSCSF--NEHGYNHFOAMRLGVEEINNST 90

Db 31 MGDIIIGALFVSVH-----HQPVDKVKHERKCGAVRQYGIQVREAMLHRLERINSDP 83

Qy 91 ALLPNTTGLYQLYDCSDSANYYATLRYLSLPGQHIEILOGLLHYS----- 137

Db 84 TLLPNTTGLCIRDSQWHS-----VALEGSIEIRSLISSEEEGLVRCVDGS 133

Qy 138 -----PTVLAVIGPDSTRRAATTAALLSPFLVPMISVYASSPTLSVKRQYSEFLRTIPN 191

Db 134 SSFRSKRPVIGYIGSSSSVAIQVONLQLENIPOIAVSATSDLSDKTLFFYFMKVVPS 193

Qy 192 DKYQVETVLLQKFGMTWISLVGSSDDYGLQVQALENOATGOGICIA--FKDIMFESA 249

Db 194 DAOQARAWVDYKRYNMVYSAVHTGNGYSGMEAFKMSAKGEGICIAHSYKITSNAGE 253

Qy 230 QVGERMOCIMHNLQAQATVYVVSRRQLARVFEESVVLNLTGK--VWVASEANALSRH 308

Db 234 GSFDFKLLKRLSRHPR--ARVYACFCEGVTYGLLMMARRLLAGFELIGSDGMADRYD 311

Qy 309 IGVGVIQRIQVNLGVALQKRAVPLKAFEEYARADKRAPRCHKGSCSSNOCLREQO 368

Db 312 VV--DGIYQ--EAVGGITTLKQS--PDVKWPDYLL--KLRPETNLRNPV-----FQ 355

Qy 369 AFMAHTMP--KLKAFSM--SSAYN-----AVYAAVAVAHGLH--OL 403

Db 356 EFWQIRFCQRLDEGFAQENSKYKTKGSSITLRTNHVQDSKMGFVNAITISMAVGLHNMOM 415

Qy 404 LGCA--SGACSR-----GRVYPMOLLQIHKVHFL--LHKDTVAFNDNRDPLSSYNIAND 455

Db 416 SLCPGYAGLCDAKMRPLDR-----KLDSLMKTNTGVSQDMILLFDENGDSPPRYEELMNRK 471

Qy 456 WNGPMTPLFVLCSSTPVQOLINETKTIQWKGKNOVPRKSVSSDCLSEGHQRYV--TGFHH 514

Db 472 EMGKDY--FDYINVGSNDNGELMKDDEV-W-SKNNNIIRSVCSBCEKQGIKVRKEYS 528

Qy 515 CCEFCPCAGTFLNKSIDLYRCOPCKEEMAEQSGOTCFPRVYVFLALREHTSWLLAAN 574

Db 529 CCWTCITPCENENY--FDREYTKACQGLSGMPDIDLGCGLIFVQYLIRMGDPERPIAAVFA 586

Qy 575 TLLLLLTGAGLIFAMHLDTPVRSAGRLCFPLMGLSLAAGSGSLYGFEPTPACILR 634

Db 587 CGLLATLAEVTVIETIYRDPVAKSSRELGYIILAGICLGLCFCLAKPKQIYCYLO 646

Qy 635 QALFALGFIIFSLCTVRSFOLLIF----KESKKVPEFHYAWONHAGLFVWISSA 689

Db 647 RIGIGLSPAMSYALVTKTNRRIARILAGSKKICTKRPMNSACAQIVIAFLICI--- 702

Qy 690 QLLICTLWVMTPLPAREYORPHL--VMECTETNSIGFLIAFLYNGLSISAFASY 747

Db 703 -QLGIIVAFIMERPDIIMDY---PSIREVYLICNTTN-LGVYTPGLYNGLLISTCFYFA 758

Qy 748 LCKDLPENYNEAKCVTFSLFNFSWIAFFTASYDGKYLPAANMAGLS-SLSSGFG 806

Db 759 EKTRNVPANFNENAKYIAFTVYTCIIMLAF---VPYFGSNYKIIIMCFSVSLSATVALG 815

Qy 807 YELPKCYVILCRPDNSTEHFQAS 830

Db 816 CMFPRVYIILAKPERNWSAFTTS 839

RESULT 9

A41939 A:Accession: A41939

A:Status: Preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-1199 <HOU>

A:Cross-references: GB:M61099; NID:9397806; PIDN:AAA19497.1; PID:g204460

A:Experimental source: cerebellum

A>Note: Sequence extracted from NCBI backbone (NCBIPI:60785)

A:Residues: 1-1199 <MAS>

A:Cross-references: EMBL:X57569; NID:956646; PIDN:CAA40799.1; PID:g56647

A:Title: Sequence and expression of a metabotropic glutamate receptor.

A:Reference number: S15362; MUID:91156047

A:Accession: S15362

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-1199 <MAS>

A:Cross-references: EMBL:X57569; NID:956646; PIDN:CAA40799.1; PID:g56647

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.2%; Score 727; DB 2; Length 1199;

Best Local Similarity 26.7%; Pred. No. 3.8e-47;

Matches 230; Conservative 155; Mismatches 362; Indels 116; Gaps 32;

Qy 34 LPGDYLLAGLFPPLHSGCLQVRRHREY-TLCDRSCSF--NEHGYNHFOAMRLGVEEINNST 90

Db 41 MGDIIIGALFVSVH-----HQPVDKVKHERKCGAVRQYGIQVREAMLHRLERINSDP 83

Qy 91 ALLPNTTGLYQLYDCSDSANYYATLRYLSLPGQHIEILOGLLHYS----- 134

Db 94 TLLPNTTGLCIRDSQWHS-----VALEGSIEIRSLISSEEEGLVRCVDGS 133

Qy 135 -----PTVLAVIGPDSTRRAATTAALLSPFLVPMISVYASSPTLSVKRQYSEFLRTIPN 191

Db 134 SSFRSKRPVIGYIGSSSSVAIQVONLQLENIPOIAVSATSDLSDKTLFFYFMKVVPS 193

Qy 192 DKYQVETVLLQKFGMTWISLVGSSDDYGLQVQALENOATGOGICIA--FKDIMFESA 249

Db 194 DAOQARAWVDYKRYNMVYSAVHTGNGYSGMEAFKMSAKGEGICIAHSYKITSNAGE 253

Qy 230 QVGERMOCIMHNLQAQATVYVVSRRQLARVFEESVVLNLTGK--VWVASEANALSRH 308

Db 234 GSFDFKLLKRLSRHPR--ARVYACFCEGVTYGLLMMARRLLAGFELIGSDGMADRYD 311

Qy 309 IGVGVIQRIQVNLGVALQKRAVPLKAFEEYARADKRAPRCHKGSCSSNOCLREQO 368

Db 312 VV--DGIYQ--EAVGGITTLKQS--PDVKWPDYLL--KLRPETNLRNPV-----FQ 355

Qy 369 AFMAHTMP--KLKAFSM--SSAYN-----AVYAAVAVAHGLH--OL 403

Db 356 EFWQIRFCQRLDEGFAQENSKYKTKGSSITLRTNHVQDSKMGFVNAITISMAVGLHNMOM 415

Qy 404 LGCA--SGACSR-----GRVYPMOLLQIHKVHFL--LHKDTVAFNDNRDPLSSYNIAND 455

Db 416 SLCPGYAGLCDAKMRPLDR-----KLDSLMKTNTGVSQDMILLFDENGDSPPRYEELMNRK 471

Qy 456 WNGPMTPLFVLCSSTPVQOLINETKTIQWKGKNOVPRKSVSSDCLSEGHQRYV--TGFHH 514

```

Db 204 VVPSDQLAARMIDIVKRYMWTYSAVHTGNTGSGMDAFKELAAQEGICIANSD--KI 261
QY 248 SAQVDERMOCAMRHAAQ--AGATVVVFSRQLAARFESVLTNLTKV--WVASEAMA 304
Db 262 YSNAGEFSPRLRLKLEERLPKARVVCFCGEMTVRGLLSMRRLGAVGSESLSGMA 321
QY 305 LSRHITVPGIQRIGMWLGAIOKRAVPGLKAPREAY--ARADEARPPCKSGSCSNQ 362
Db 322 DRDEVEIEGYEANG--GTTIKLOS--PEVRSFDDYFLKRLDNTNPNPPEFWQHRFQ 377
QY 363 LCBEQAQFMAHTM--PKLKAF-----SMSAY-----NAYRAVYAVANGL-----HQLL 404
Db 378 -CR-----LPGHLENPFKVCYCNESLEEVYQDSKMGFYNAITYAMHQLQMHMHALC 432
QY 405 GCASGACSR-----GRVYPMOLLEQIHKVHFL-LHKDTVAFNDRDPLSSYNTIAMD-N 457
Db 433 PGHVGICDAMKPIDGR-----KLDFLIKSFVGVSGEEVWFDEKGDAPGRYDINLQYTE 488
QY 458 GPKWTFVVLSSWSPVQVLNINETKIQMHKDNQPKSVSCSDGLEHQRVY--TGFHHC 516
Db 489 ANRDYHVAG--TWHEGVNLIDDKIQM--KSGMV-RSVSEPCLKQIVIRKGEVSCC 544
QY 517 FECVPCGAGTFLNKSIDLRCOPCKEEMAEBSQTCFPRVVFALREHTSWVLLAANTL 576
Db 545 WICACAKENEFV--QDEFICRACDLGMMRNALGCEPIPRVRYLEMDSIIAIAFSL 602
QY 577 LLLLLGTAQLFAMHLDTPVYRSGRLCFLMGLSAGSGSLGFGPEPTPACLLRQA 636
Db 603 GIATVTEVTLIFLYRPTPVKSSRELQYIIAGIFLGCPPTLAKPTTSCYLRL 662
QY 637 LPAFGTIFLSCILVRSEOLIIF-----KFSFKVPTFYHAMVONHAGLEFVMTSS--A 688
Db 663 LVGLSSMCSALVTKNRRLARILAGSKKICTRKPRMSMAQ-----VIIASLIIS 715
QY 689 AOLLICTMLVWVTPLEPAREYORPHLMCEETNSLGFLEFLVNLGLSISAFASYL 748
Db 716 VOELTVTLTIMERPMDILSPISKEVYLL--CNMSNLGVAVAPGVYNGLLIMSTCYAFK 773
QY 749 GKDLPEYNNAKCVTFESLLENFVSWIAFPTTASYDG--KILPAAANMAGSLSSGGGY 807
Db 774 TRVNPANFNNAKYIAFTMYTCIIMIAF-----VPIYFGSNYKIITTCFVAVSLVVALGCM 830
QY 808 FLPCYVILCRPDLNSTEHFOAS 830
Db 831 FTPKMTIILAKPRKNVSAFTTS 853

RESULT 10
JH0563
metabotropic glutamate receptor 4 precursor - rat
C:Species: Rattus norvegicus (Norway Rat)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C:Accession: JH0563; 158149
R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A:title: A family of metabotropic glutamate receptors.
A:Reference number: JH0561; M0ID:92110002
A:Accession: JH0563
A:Molecule type: mRNA
A:Residues: 1-912 <TRAN>
A:Experimental source: brain
R:O'Hara, P.J.; Shepard, P.O.; Thøgersen, H.; Venezia, D.; Haldeman, B.A.; McGrane, V.;
Neuron 11, 41-52, 1993
A:title: The ligand-binding domain in metabotropic glutamate receptors is related to bac
A:Reference number: 158149; M0ID:93332699
A:Accession: 158149
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-123, 'R', 125-912 <RES>
A:Cross-references: GB:M05018; NID:g205400; PIDN:AAA93190.1; PID:g205401
C:Comment: This protein is coupled to a G protein and evokes a variety of functions by m
C:Genetics:

```

```

A:Gene: GLUR4
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phos
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-912/Product: metabotropic glutamate receptor 4 #status predicted <MET>
F:568-912/Domain: transmembrane #status predicted <TR>
F:625-645/Domain: transmembrane #status predicted <TR>
F:657-675/Domain: transmembrane #status predicted <TR>
F:700-720/Domain: transmembrane #status predicted <TR>
F:751-772/Domain: transmembrane #status predicted <TR>
F:786-807/Domain: transmembrane #status predicted <TR>
F:822-847/Domain: transmembrane #status predicted <TR>
F:98,301,454,484,569/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:621,689,695,859,870/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 15.8%; Score 708.5; DB 2; Length 912;
Best Local Similarity 26.6%; Pred. No. 7e-46;
Matches 227; Conservative 145; Mismatches 371; Indels 111; Gaps 32;

QY 34 LPCDYLLAGFLPLH-----SGCQVRRHREPVTLCDNSCFNENGYLFOAMRLGVEIN 87
Db 45 IDGDITLGLFPPVHGRSEGRKACGELK-----KEKGIRHLEAMFLPALBRIN 90
QY 88 NSTALLPNITGLQYQVDCSDSANVYATRLYSLPGQHHELOG-----DLIHY 137
Db 91 NDPDLPNITLIGARILDTLC--SRDTHALEOSTLV--QALIKDGETEVRCGSGGPIITKP 147
QY 138 PTVLAVIPGSTNRAATTAALLSPFLVPMIYAASETLSYKRVPSFLRTIPDKYQVE 197
Db 148 ERVYGVIGASGVSSIVMANILRFKIPQISYASTARDLSNRSYRDFSRVPSDYOAO 207
QY 198 TMVLLDQFGWTWISLVGSSSDYQOLGYQ--LEMQATGQIGICAFKDIIMPSSAOGBERM 256
Db 208 AMVDIVRALKMNWYVSTLASSEGSVEAEIQKSRRENGVCIASVYKIPREPRTG--EPD 266
QY 257 QCLMRHHAQAGATVYVVFSSRQLAARFESVLTNLTKV--WVASEAMALSRHTTYPGI 315
Db 267 KIIRLETSNARKIITIFANEDDIRVLEAARANQGHFFMMSDSMGSS--APVRL 324
QY 316 QRIQWLVGVALIQRAVPGALKAFAEYAVRADKAPRCHKSGWS-----SNQCREQAFM 371
Db 325 EEVAGAVTILPKRM--SVRGFDRTYFSSRLDNNR-----RNIMFAEFMEDNHCK-----LS 375
QY 372 AHMPK-----LAFSSMSAYN-----AYRAYANAHGLH-----QLLGASAGC 411
Db 376 RHALKKSHIKKCTNRRIIGDSAYEQEGKQVYIDAVYAGHALHMHRLDLCGRVGLC 435
QY 412 SR--GRVYPMOLLEQIHKVHFL-LHKDTVAFNDRDPLSSYNTIAMDNGPKWTFVLGSS 469
Db 436 PRMDPVGTQOLKTYIRVNFSGINGNVTTFENGDAGRDIYQOLRNGSAEKVIGS- 494
QY 470 TWSPVOLININETKIQMHKDNQPKSVSCSDGLEHQRVYTFPHHCPEVCYPCGAGTFPLN 529
Db 495 -WTD-HILRLIERMQMGSGQOLPRISICSLPCQGERKKYKVMACCMHCPCPTG--YQY 550
QY 530 KSDLYRQOPCKEEMAEBSQTCFPRVVFALREHTSWVLLAANTLTLTLTLTAG-- 586
Db 551 QVDRYTKCTCPYDMRPLENRTSCQPIPV--KLEWDSFWALVP---LFAVVGIAATL 604
QY 587 ---LFAHMLDTPVYRSGRLCFLMGLSAGSGSLGFGPEPTPACLLRQALFALGFT 643
Db 605 VVVFVYRNTDPIYKASGRRELSYLLAGIFLCYATTTLMIAEPDLGRCSLRITLGLGMS 664
QY 644 IFLSCLTVRSFOLIIF--KFSKVPFEYHAWQNHAGLFWISSAOLILCTLWLV 700
Db 665 ISYALTLKTNRIYRIFEOGKRYSAPRFLSPASQL--AIFILIS--LQLLGICVAFVV 720
QY 701 WTPLEPAREYQ-----RPHVLMLCEETNSLGFLEFLVNLGLSISAFACSLGLDLP 753
Db 721 DPHSVVDVFOQRLDPLDFARGV--LKC-DISDLSLCLLGLSMLLMWTCYVAYAKTKGVP 778
QY 754 ENYNACQVTFSLDFNEVSWIA-----FTTPASVYDKYLLPAAANMAGLS--SLSSGFGGYF 808

```



```

Db 137 TGVIGSYSDVSIQVANLLRFLQIPQISYASTSAKLSDKSRDYFARTVPDPFOAKAMA 196
Qy 201 LLLQKFGMTWISLVGSSDDYQGLQVQALQENATGOGICIAKDIIMPSSAQVGDERMOCIM 260
Db 197 EILFFEMWTYVSTVASBDDYGETIEAFELARARNICVA-----TSEKGRAMSRAAF 250
Qy 261 RHLQAQ-----GATVVVVFSSRQLARVFEEFESVLTNLTKGVASEAMALSRHITGVGCI 315
Db 251 EGVVRLALQKRSARVAVLFTFSEDARELLATQRLN-ASFVWVASDQMGALSFV----- 303
Qy 316 QRIQMVGLVALQKRAVPGKAEFEA-YARAD-----KRAPRCHKGSW-----C 358
Db 304 -----VAGSERAAAGAITIELASYPIPSASYFQSIDMNNSSRNMPWREFEERPHC 355
Qy 359 SSNQLCRECAFMAMHTPKLAKAFSSMSAYNAYRAVYAVAHGLHQLG-----CASG 409
Db 356 SFPR--KDCAAHSLRAVFEESKIMFVNV--AVYAAHALLHMHRLALCPNTHLCLDAM 410
Qy 410 ACSGRVYPMQLERQIKHVEHLH-----KDTVAFNDNRPLSSYNTIAMDMN 457
Db 411 RPNVGR-----RLYK-DEVLVNKKFDAPRPADTDDEVDFRFGDQIGKYNIFTYARA 461
Qy 458 GP-KWTFYVGLSSVSPVQVQINETKIQWNGKD-NQVPKSVSSDCLGHRVVTGFPHC 515
Db 462 GSGRYRQKVG--YMAE-GTLDTSFIPWASPSAGPLPASCSEPCLONEVKSVPQGEVC 518
Qy 516 CFECVPCGAGTFLNKSQDYRCQCKEMAPESQF--CFRTVVFALREHTSVL--LA 572
Db 519 CMLDTPQOPERY--RLDEFQADCGLGW--PNASLTGCFELPQEF--TRMGDAMAVGPT 573
Qy 573 ANTLLELLLLTAGLTFAMHLDTPVRSAGRLCFMLGSLAAGSGSLYGFGEPTPACL 632
Db 574 IACGALATLEVLQVFRHNAFTPVKASGRELCTILGVLGVCMTFVFAKSTANCT 633
Qy 633 LRQALFALGFTIFLSCITVNSFOLITF--KSTKYPTFTHANVQNHGAGLFWISSAA 689
Db 634 LRRIGLTAFSVCSALLTKTNRIARIFGARREGAQRPF-----ISPAS 678
Qy 690 QLLTCL-----TWLVVWTP-----LPAREYGRPHLVMLBCT--ETNSLIGFIL 730
Db 679 QVALCCLALISQGLIIVAMLVVEARPGKETAPER--EVTYTLCKNHHDASMLG-- 730
Qy 731 AFLVNGLLISAFACSYLGKDLPENYNEAKCVTFSLFNFVSIATFTTASVYDGKYLPA 790
Db 731 SLAVNVLILACTLYAFKTRKCPENFNAKFIQFTMYTTCILWLAFLPIFYVTSDDYVQ 790
Qy 791 ANMAGLSSLSG--FGGYFLPKCYVILCRPDLNSTEHFQASIDYTRGCS 840
Db 791 TTTMCVSVLSGVSVALGCLFAPKHLIILFQPKNVVSH-----RAPTSRFGS 837

RESULT 13
JH0562
metabotropic glutamate receptor 3 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1992 #sequence-revision 30-Jun-1992 #text-change 13-Sep-1998
C:Accession: JH0562
R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A:Title: A family of metabotropic glutamate receptors.
A:Reference number: JH0561; MUID:92110002
A:Accession: JH0562
A:Molecule type: mRNA
A:Residues: 1-879 <TAN>
A:Experimental source: Brain
C:Comment: This protein is coupled to a G protein and evokes a variety of functions by m
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F:1-22/Domain: signal sequence #status predicted <sig>
F:23-879/Product: metabotropic glutamate receptor 3 #status predicted <MET>
F:377-599/Domain: transmembrane #status predicted <TRI>
F:614-634/Domain: transmembrane #status predicted <TII>
F:646-664/Domain: transmembrane #status predicted <III>

```

```

F:689-709/Domain: transmembrane #status predicted <TIY>
F:735-756/Domain: transmembrane #status predicted <TRV>
F:770-791/Domain: transmembrane #status predicted <TVI>
F:804-828/Domain: transmembrane #status predicted <VLI>
F:209,292,414,439/Binding site: carbohydrate (asn) (covalent) #status predicted
F:610,845/Binding site: phosphate (ser) (covalent) #status predicted

```

Query Match 15.3%; Score 688; DB 2; Length 879;

Best Local Similarity 25.8%; Pred. No. 2,56-44;

Matches 227; Conservative 139; Mismatches 355; Indels 160; Gaps 35;

```

Qy 31 DFTLPQDYLGLTLPRLHSGQLQVRHREPVTLCDRSCSFNE-HGYHLFOAMRLGVEEINNS 89
Db 32 EIKIEGLDVLGLPINE-----KGTETECGRINERGRIGRLQLEAMFALDEINKD 82
Qy 90 TALLPNTLGYQLYDVCSANVYATRLVSL-----PGQHIELOGLD 133
Db 83 NYLLPGVYKLVGHILDTG--SRDTYALBQSLFVFRASLTQVDEAEYMCPSGAYAIQENIPL 140
Qy 134 LHYSPYLAIVGPDSTNAATTAALLSPFLVPMISYASSSETISYKRPYFELTIPMDK 193
Db 141 L-----IAGVIGGSYSYSIQVANLLRFLQIPQISYASTSAKLSDKSYDYFARTVPDPF 195
Qy 194 YQVETVNLLOKFGMTWISLVGSSDDYQGLQVQALQENATGOGICIAFKDIMPSAQVGD 253
Db 196 YQAKAMAILFFMWTYVSTVASBDDYGETIEAFEGARLRARNICIAFAE-----KVG 249
Qy 254 ERMQ-----CLMRHLAQ-AGATVVVVFSSRQLARVFEEFESVLTNLTKGVASEAMALSRH 308
Db 250 SNIRKSYDVIRELLQKPNARVVVLFMRSDSRRLIAANRVN-ASFVWVASDQMGAOES 308
Qy 309 ITGPGIQRIGMVGVALQKRAVPGKAEFEAYARADAEARPOHKSQW-----C 358
Db 309 I--VKGSEHV--AVGATLELAHPVROFDRFQSLN--PYNNRKPFEDFENQKFOC 361
Qy 359 S-----SNQLCRECAFMAMHTPKLAKAFSSMSAYNAYRAVYAVAHGLHQ-----LLGAS 408
Db 362 SLQKRNHRQYCKHIALIDSNEYQ-----ESKIMFVNVAYAMAHKMQRLCPRTT 416
Qy 409 GACSRGRVYPMQL--EQIHVVEHLHKDYAFENDR-----DPLSSNTIAW 454
Db 417 KLCDAKMLIDKKLYKELKINF-----TAPFNKKGADSIKVFDPDGMGRVNVNL 471
Qy 455 DMNPGKMTFYLGSSVSPVQVQINETKIQWNGKDQVPKSVSSDCLGHRVVTGFPHH 514
Db 472 QQTGKYSYLKVGH--WAEF--LSLDVDSIHV--SRNSVPTSCSDPCAPNEMKNMQPQDV 526
Qy 515 CFECVPCGAGTFLNKSQDYRCQCKEMAPESQF--PRVVFALREHTSW----- 568
Db 527 CCWICIRPEPEYEL--VDEFQCMCGPGQWPTADLSGYNLPEDYI-----KMEAMALGP 580
Qy 569 VLLAANTLLLLTAGLTFAMHLDTPVRSAGRLCFMLGSLAAGSGSLYGFGEPTR 628
Db 581 VTIAICIGLTCIYIT--VFYKHNHTPLVKASGRELCTILGVSLSYCMTEFFFAKSP 638
Qy 629 PACILRQALFALGFTIFLSCITVNSFOLITF--KSTKYPTFTHANVQNHGAGLFWMI 685
Db 639 VICARIRGLCTSPATCYSAITLTNCIARLFDGVKCAQQRKF-----I 683
Qy 686 SSAADLLTCL-----TWLVVWTP-----LPAREYGRPHLVMLBCTETNSIGF 728
Db 684 SPSSGVFICLDLIIQVIMVSMVILDERPGRRTLLPEKR-----EYILKCNWKS-SM 737
Qy 729 ILAFLVNGLLISAFACSYLGKDLPENYNEAKCVTFSLFNFVSIATFTTASVYDGKYL 788
Db 738 LISLYDVAVLITCTVYAFKTRKCPENFNAKFIQFTMYTTCILWLAFLPIFYVTSDDY 797
Qy 789 PAANMAGLSSLSG--GGYFLPKCYVILCRPDLNSTEH 826
Db 798 VQTTMC--ISVLSGVSVALGCLFAPKHLIILFQPKNVVTH 837

```

RESULT 14
JC7160
metabotropic glutamate receptor subtype 3 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 11-May-2000
R:Accession: JC7160
R:Minoshima, T.; Nakanishi, S.
J. Biochem. 126, 889-896, 1999
A:Title: Structural organization of the mouse metabotropic glutamate receptor subtype 3
A:Reference number: JC7160; MUID:20012997
A:Accession: JC7160
A:Molecule type: DNA
A:Residues: 1-879 <MIN>
A:Cross-references: GB:AF170696
C:Genetics:
A:Gene: mGluR3
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: differentiation; G protein-coupled receptor; receptor; transmembrane protein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-879/Product: metabotropic glutamate receptor subtype 3 #status predicted <MAT>
F:577-599/Domain: transmembrane #status predicted <TM1>
F:614-634/Domain: transmembrane #status predicted <TM2>
F:646-664/Domain: transmembrane #status predicted <TM3>
F:688-709/Domain: transmembrane #status predicted <TM4>
F:735-756/Domain: transmembrane #status predicted <TM5>
F:770-791/Domain: transmembrane #status predicted <TM6>
F:804-828/Domain: transmembrane #status predicted <TM7>

Query Match 15.3%; Score 688; DB 2; Length 879;
Best Local Similarity 25.9%; Pred. No. 2.5e-44;
Matches 229; Conservative 142; Mismatches 346; Indels 168; Gaps 36;

QY 31 DFTLEGDYLLAGLPPLH--SGCLQVRHREYTLCDRSCSFENHGVHLFOAMRLGEEIN 87
DB 32 EKIKGSDLVVGLFPINKEGTETEECKINE-----DRGIQRLKLAFLPAIDELN 80
QY 88 NSTALLPNTITGYOYDVCSDSANVATLRLVSL-----PCQHHIELOG 131
DB 81 KQNYLLPGVKLGVLHIDTC--SRDTYALEQSLFVRASLTKVDEAEVYMCPOGSVAIQENI 138
QY 132 DLLHSPYLVAVIGDSTNRKATTAALLSPFLVPMISVASETISVKQYPSFLRTIN 191
DB 139 PLL-----TAGVIGSYSSVSIOVANLRLFIQIPISTASTAKSDSKRYDYFARTYVP 193
QY 192 DKYQVETWVLLQKFGWTISLVGSSDDYGQVQVALENQATGOGICIFKIDIMFSAOV 251
DB 194 DRYQAKAAEILRYNWTYVTSVASEGDYGEICIAFEDEALRNICITAE-----KV 247
QY 252 GDERMO---CLMRHQAQ-AGATVVVF---SSROLARVFESVYLTNLGKVVAVASEA 302
DB 248 GRSNIRKSYDSYIRBLQKPNARVYVLFWRSDSREL--IAASRVNASFT---VWASDG 302
QY 303 WALSHHTGVPQIQIIGWLVGIAIOKRAVPGAKAFEEAARADKFAPRCHGSGW----- 357
DB 303 WQAQEST--VKQSEHV--AYGAITLLEASHPVROPDRYFQSLN---PYNHNNHFWRFDEW 355
QY 358 -----CS-----SNQLCRECOAFMAHTMPKLAFSMSASVAVYAVVAHGLHO----- 402
DB 356 EOKPQCSLQNKRNHRQICDKHLAIDSSNEYQ-----ESTIMVAVNAVYAMALHKKMOT 410
QY 403 LIGCASGACSRGRVYPMOILQIHKVHFLHKDVAFNDNR-----DPLSSYN 450
DB 411 LCPNTTKLDAKRIIDG---KRLKYDYLLKINFAPFNPKGADSIYKFDYGDGGRYN 467
QY 451 IIAAMNMPKMTFYVIGSSTWSPVOLINETKIOMHGKNOVPKSYCSSDCLEGRQRYVT 510
DB 468 VFNFIHIGKSYLVKVGH--WAET--LYLDVDSIHW--SRNSVPTSQCDSPCAPNEKKMNO 522
QY 511 GHNHCCEFCVPCGAGTFLNKSIDLYRCOPCGKEWAPESQOTCF--PRVVFALRPHSTW 568
DB 523 PDQVCCWICIRPEPYEYL--VDEFTCMDCGPGQWPTADLSGCTINLPEDYI-----KEDAM 576

QY 569 ----VLLAANTLLLLLTGTAFLFAMHLDTPVRSAGRLCFMLGSLAAGSGSLYFEG 624
DB 577 AIGPVIACLGFMCIQIVT--VFIKHNNTPLVKASGRLECYLLFGLSYSCMCFEFLA 634
QY 625 EPTFRACILROLALFALGFIFLISCLTVSFOIITF---KFTKPTTYHANVQNHAGL 681
DB 635 KPSVICALRRLIGTSTFAICYSALLTNTCIARLEFDGVKNAQRPKF----- 682
QY 682 FWMISAAQLLCL-----TWLVWMP-----LPAREYRPPHVLMECTEEN 724
DB 683 ----ISPSSQVFLCCLLIVQIYVMSVWMLLEPGRRTLPKRR-----EYIILKCNKD 734
QY 725 SLGFIILFVNGLLSISAFACSYLCKLDPENYNAKCVTFSLLENFVSIAPFTTASYD 784
DB 735 S-SMLISLTVDVVILCTVAFKTRKCPDENNEAKFIFGTWYTCIIMLAFLPFYVTS 793
QY 785 KYLLPRAANMAGLSLSGSGF---GGYFLPKCVILLCRPDLNTEH 826
DB 794 SDYRQVTTMC-IVSLSGFVVLGCLFAPKRVHIVLFQPKNVVTH 837

RESULT 15
A46742
metabotropic glutamate receptor, mGluR6 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
R:Accession: A46742
R:Nakajima, Y.; Iwakabe, H.; Akazawa, C.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakanishi, T.
J. Biol. Chem. 268, 11868-11873, 1993
A:Title: Molecular characterization of a novel retinal metabotropic glutamate receptor
A:Reference number: A46742; MUID:93280152
A:Accession: A46742
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-871 <NAK>
A:Cross-references: GB:D13963; NID:q391856; PIDN:BA003066.1; PID:q391857
A:Experimental source: retina
A:Note: sequence extracted from NCBI backbone (NCBIN:133246, NCBIPI:133250)
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 14.8%; Score 664; DB 2; Length 871;
Best Local Similarity 26.2%; Pred. No. 1.7e-42;
Matches 236; Conservative 134; Mismatches 378; Indels 152; Gaps 35;

QY 13 ILISCCW-----AFACHSSESPDFTLPQDYLLAGLPPLHSGCLQVRHREYTLCDRSCS 67
DB 7 LLLMLAMWLSQAGIAC---GAGSVRLAGGLTGLGFLPVHAR-----GAAGRCAG 52
QY 68 F--NEHGYHLFQAMRLGYEEINNSTALLPNTITGYQYDVCSDSANVATLRLVSLPQOH 125
DB 53 ALKKQGVHRLKAMLYALDRVNADELLPGVRLGARLLDTC--SRDTYALEQSLSFV--QA 109
QY 126 HIELOGD-----LLHISP--TVLAVIPDSTNRATTAALLSPFLVPMISA 170
DB 110 LIRGRDGDDEASVRCPGVPLRSAPPRVAVAVASASVSIMAVNLRFAIPIQISA 169
QY 171 ASSETLSVRYKQVPSFLRTIPNDKYOVETWVLLQKFGWTISLVGSSDDYGQVQVALE 229
DB 170 STAPELSDSTRIDFSSRVVPPDSIOQAMVDIVRLGMNHYSTLASEGNGESGVEAYVQ 229
QY 230 NQATGQICIAFKDIMPSPSAQVDERMOCMLRHLAQAATVVVYSSROLARVFESVYL 289
DB 230 ISREAGVCIASIKIIPREPKRG--EFHVIYIRIMETPNARGLIIFANDEDIRRVLEARTQ 288
QY 290 TNLTKG--VWVASEANALSRHITGVPQIQRIGWLVGIAIOKRAVPGAKAFEEAARADKEA 348
DB 289 ANLTGHFLVWGSWSW--GSKISPIINLEEAIVGATIIIPKRA--SIDCFDQYFTRSLEN 344
QY 349 PRPCHKGSW-----C-----SSNQLCRECOAFMAHTMPKLAFSMSASVAVN--- 388
DB 345 NR--RNIMFAEFWEENFNCKLITSSGOSDSDSTRCTG-----EERIGODSAVEEDG 393

```

OY 389 -----AYRAVYVANGHQ-----LLGCSAGC-----SRGVYPMOLLEQIHNKXPHLLHKD 435
Dh 394 KQVQYIDAVYAIINALISMSHOALCPGHTGICLCPAMEPIDGR-----LHATRAVAFNSAG 449
OY 436 T-VAENDNRDELSSYNIAND-WNGPKWTFYVIGSSYWSFVQVQININTKIOMHGKDQVP 493
Dh 450 TPVAMENENGADPARGYDIFQYQATNGSSASSGGYQAVQOWAEA-LRLDEYVLRMSGDPPHEVP 508
OY 494 KSVSSDCLBEHORVYVGHHCCEYCPCCAGFTPLKNSDLYRQPCCKEEMAEBSQYCF 553
Dh 509 PQSCSLPGCPERRKKMKYGPCCHNCACOGYR--OVDEFTCEACGDMRPTPNHTGR 566
OY 554 PRTVYALAREHTSMVLIANTLLILLILG-----TAGFEAMHLDPPRVASAGRLCFL 607
Dh 567 PTVYVRL-----TWSSPMALPELLAVLGIMATTIMAFEMRNDRPIYRASGRELSTY 620
OY 608 MUGSLAAGSGSLYCFEGEPTRPACLLQALFALGFYFLSCLYVRSQULIIF---KFST 664
Dh 621 LLTGIFLIYATFTFMAVEPCAIIOAARLLLGIGTTLSTYALLLTKNIRYIEQGRSV 680
OY 665 KVPFTFYHAMVQNHGAGLFVMISSAQAOLI-----CLTWLVWYPTLPRREYOR-- 711
Dh 681 TTPPF-----ISPTSQLVITFGLTSLQVGVIMLGAQPHSHVIDEQR 725
OY 712 --PPhLV-MIECTETWSLGLIAFLAYLNGLSISAFACSTYLGKDLPENYNEAKCVTFSL 767
Dh 726 TVDEQOARGVLYKC-DMSDLSLTIGLGISLLIMWYCYVATAKAGVPETFEAKPIDGTYM 784
OY 768 FNFPSMIA----PFTTASVYDQKFLPAAANNMAGLS--SLSGSGFYGLPKCYVILCRPDLN 822
Dh 785 TTCTIIMIAFVIEFGTQSOAKETIYQTTTLTVLSLSLASVSLGMLVYPKTYVILFHEQN 844

```

Search completed: June 11, 2002, 10:19:35
Job time: 2829 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 09:19:53 ; Search time 24.84 Seconds
(without alignments)
826.970 Million cell updates/sec

Title: US-09-819-946-2
Perfect score: 4493
Sequence: 1 MLCTARLVGLLISCCWA.....NSTEHQASIDYTRCGST 841

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
Issued_Patents_AA:*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1142	25.4	1078	1 US-08-485-588-7	Sequence 7, Appl1
2	1142	25.4	1078	1 US-08-484-565-7	Sequence 7, Appl1
3	1142	25.4	1078	2 US-08-480-751-7	Sequence 7, Appl1
4	1142	25.4	1078	2 US-08-943-986-7	Sequence 7, Appl1
5	1142	25.4	1078	3 US-08-353-784-7	Sequence 7, Appl1
6	1142	25.4	1078	3 US-08-484-719B-7	Sequence 7, Appl1
7	1142	25.4	1078	4 US-08-484-159-7	Sequence 7, Appl1
8	1130	25.2	1088	1 US-08-485-588-6	Sequence 6, Appl1
9	1130	25.2	1088	1 US-08-484-565-6	Sequence 6, Appl1
10	1130	25.2	1088	2 US-08-480-751-6	Sequence 6, Appl1
11	1130	25.2	1088	2 US-08-943-986-6	Sequence 6, Appl1
12	1130	25.2	1088	3 US-08-353-784-6	Sequence 6, Appl1
13	1130	25.2	1088	3 US-08-484-719B-6	Sequence 6, Appl1
14	1130	25.2	1088	4 US-08-484-159-6	Sequence 6, Appl1
15	1127.5	25.1	1085	1 US-08-485-588-5	Sequence 5, Appl1
16	1126.5	25.1	1085	1 US-08-484-565-5	Sequence 5, Appl1
17	1126.5	25.1	1085	2 US-08-480-751-5	Sequence 5, Appl1
18	1126.5	25.1	1085	2 US-08-943-986-5	Sequence 5, Appl1
19	1126.5	25.1	1085	3 US-08-353-784-5	Sequence 5, Appl1
20	1126.5	25.1	1085	3 US-08-484-719B-5	Sequence 5, Appl1
21	1126.5	25.1	1085	4 US-08-484-159-5	Sequence 5, Appl1
22	1126.5	25.1	1085	4 US-08-485-588-8	Sequence 8, Appl1
23	1123	25.0	1079	1 US-08-484-565-8	Sequence 8, Appl1
24	1123	25.0	1079	2 US-08-480-751-8	Sequence 8, Appl1
25	1123	25.0	1079	2 US-08-943-986-8	Sequence 8, Appl1
26	1123	25.0	1079	3 US-08-353-784-8	Sequence 8, Appl1
27	1123	25.0	1079	3 US-08-484-719B-8	Sequence 8, Appl1

28	1123	25.0	1079	3 US-08-484-719B-8	Sequence 8, Appl1
29	1123	25.0	1079	4 US-08-484-159-8	Sequence 8, Appl1
30	1093.5	24.3	1027	4 US-09-162-021B-2	Sequence 2, Appl1
31	1088	24.2	1219	2 US-08-687-289A-6	Sequence 6, Appl1
32	777	17.3	1058	2 US-08-687-289A-5	Sequence 5, Appl1
33	741.5	16.5	1180	4 US-08-660-148-2	Sequence 2, Appl1
34	741.5	16.5	1212	4 US-08-660-148-5	Sequence 5, Appl1
35	733	16.3	1056	2 US-08-687-289A-7	Sequence 7, Appl1
36	727	16.2	906	5 PCT-US91-09422-17	Sequence 17, Appl1
37	727	16.2	1199	1 US-08-041-338-2	Sequence 2, Appl1
38	727	16.2	1199	1 US-08-463-642-2	Sequence 2, Appl1
39	727	16.2	1199	1 US-08-455-602-2	Sequence 2, Appl1
40	727	16.2	1199	2 US-08-465-157-2	Sequence 2, Appl1
41	727	16.2	1199	5 PCT-US91-09422-2	Sequence 2, Appl1
42	724	16.1	906	1 US-08-486-270-2	Sequence 2, Appl1
43	724	16.1	906	3 US-08-367-264-2	Sequence 2, Appl1
44	724	16.1	1056	2 US-08-687-289A-8	Sequence 8, Appl1
45	720.5	16.0	1180	1 US-08-486-270-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-08-485-588-7
Sequence 7, Application US/08485588
Patent No. 5688938
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,588
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451

Qy	67	SENENGYLLPAMRAGVEYENNNPALLPNITLGOLYDVSQSDSNVAVATRLVSLPGOH	126
Db	63	-YNRGFRWLOAMFALFEINISSPALLPNITLGRIDFTCNYSK--ALPATISFAONK	119
Qy	127	TELQ-----DLHYSPTVIAIGPDSFNRAATTAALLSPFLVPMISAASSETLSYKRO	181
Qy	182	YPSFLRTIPDNKYOVEVPMVLLQKEGWTWLSLVQSSSDQYDOLQVQALENOATOGCTIAF	241
Db	180	FKSLRTIPDNHQTAMADITLFFRNWVGTTIADDDYGPGLEKREBAEEDDITDF	229
Qy	242	KDIMEFQAVDERMOCILMRHQAQATVVVWSSRQIARFEFESVLTMLTKWVWASE	301
Db	240	SELI--SQYSDDEEIQHVEVYQNSTAKVIYVFSQSDPLEPLKEIYRNITOKIMLASE	297
Qy	302	AMALSRHTTGPQIGORIGMWLGVAIQKRAVPGKAF-----EATARADKEAPR	351
Db	298	AMASSLIAMPQYFRHVVGTTIGFALKAGQIYGFEEFLKVVHPRSVHNGAKKEFWETFN	357
Qy	352	CH-----KG-----SWCSSNOLQRECOAFMAHMTKIKAFSSSA	386
Db	358	CHLOGAKGPIPVDTPLRGHEESDRSNSSTARPLCTODENISVETPYIDYTHLRIS	417
Qy	387	YNATRAYAAVAGHLQLGC-----ASGACSR-GNVTYMOQLLEQIHKYHFTLHK-DIV	437
Db	418	YNVLYAVYSIHAALODIYTCTCPRGLEFTNGSCADIKKVEAMOVULKHLRHNFNNMEQV	477
Qy	438	AFNNRBDLSYNIIAMDMNPKMTFFVLSSWTSPV-----OLNINETKIOH	486
Db	478	TFDEGDLVGNRTIINMHLSPED-----GSYFKEVGYNNYAKAKGERLFINDEKTIWS	531
Qy	487	GKDNQVPKSVCCSDCLEBGRH-VVTGFHCCFCECPVPCAGTFLNKSDLYRCOPGRKEWA	545
Db	532	GFSEFVPSNCSRDLCTAGTRKGIIEGEPCTCFCEVCECPDESEPTDASACNKCPODFWS	591
Qy	546	PEGQTCPRRVVLAALREHNSWVLLAANULLLLLLGTAGLFAMHLDTPVWSAGRLC	605
Db	552	NENHTSCIAKIEFLSWTEPGIALTLFAYVIGJTLFVGLVFKEFRNTPIYKATNRELS	651
Qy	606	FLMLGSLAAGSGSLYGFGEFETPRACILROALFALGFTIFLSCVTFSPOLIHFKFSTK	665
Db	652	YLLFSLTLCRSSSLDFEIGEPQDWTCLRQAPFISVPLCSITVNTNVLVFE--AK	709
Qy	666	VPTFLH-AWQNHGAGLFVMTSSAQLLICLTMLVVTPLPBARFYQFPHVLMLECHETN	724
Db	710	IPTSHRKRWGMLINQFLVLFCTMQVQIVCYIMLYTAPSPSYRQLEDEBIFFTCHGS	769
Qy	725	-SLGFIILAFYNGHLISAFACSYLICKDLPENYNEAKCYFSLFNFWMIAFETT-AS	781
Db	770	LMALGFLIG-YTCLILAACIFFAFKRSKLPENNEAKFTIFSMLIFIYMWISPIYAS	827
Qy	782	YVDGTYLPAANMAGSLSSGFCGYFLPKCYVLRCPLDNSTBHQAS	830
Db	828	TY-GKFSVAVEVIAIILASFGILACIFENKIYIILFKPSRNTIEEVRCSS	875
RESULT 4			
US-08-943-986-7			
Sequence 7, Application US/08943986			
Patent No. 5962314			
GENERAL INFORMATION:			
APPLICANT: Edward M. Brown			
APPLICANT: Steven C. Hebert			
APPLICANT: James E. Garrett, Jr.			
TITLE OF INVENTION: CALCULUM RECEPTOR-ACTIVE			
NUMBER OF SEQUENCES: 20			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Lyon & Lyon			
STREET: First Interstate World Center			
STREET: Suite 4700			
STREET: 633 West Fifth Street			

```

CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943.986
FILING DATE: 03-OCT-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/484,565
FILING DATE: 7-June-1995
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/993,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-986-7

Query Match 25.4% Score 1142; DB 2; Length 1078;
Best Local Similarity 31.0%; Fred. No. 2e-105;
Matches 276; Conservative 163; Mismatches 358; Indels 92; Gaps 22.

16 SCGW---AFACSTSSPD--FTLEGDYLLAGFLPHSGCL---QVRHREPVLTCDRSC 66
||||| : : : : || : : ||| : : : : : : : : : : : : : : : :
5 SCGWLLALHTWTSAYGDPQRAKKGDITLGLFPIHGVAAKKDDKSRPSVEICIR-- 62

67 SFNEGHYHFOAMRLGVEEINNSTALLPNTITGYOLYDVCSDSANVYATLRYLSLPGOH 126
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
63 -YNEGFMLQAMIAIEEINSSPALLPLTLGRIPTDCNTVSK--ALEALISFVAONK 119

127 IELGQ-----DLHNSPYLVAVIGDSTNRATATTAALLSPFLVPMISVAASETISVRQ 181
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
120 IDSLNLDFFCNCSEHPISTIAVAGTSGSVSTAIVANLGLFTIPQVSTASSSRLLSNKKQ 179

182 YPSFLRTIPNDKYQVETFWLLLOKFGWTYISLVGSSDDYGGQLGYVALENOATGOGICIAF 241
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 FKSFRTIIPNDHQATAMADITETFRMWNVGTIAADDDYGRGICIKFREAEERIDCIDEF 239

242 KDIMEFSAQVGDERMOCIMRHLAQAGATVVVVFSSRQLARVFESVVLNTLTGKYWVASE 301
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 240 SELI--SOYDDEEIOHVEVIONSTAKIVYFSSGPDLEPLKEIVRNITGKIMLASE 297
OY 302 AMALSRHITGVPGIORIGMVAVAIORAVPGLKAF-----EAYARADKEARPP 351
Db 298 AMASSSLIAMQYFHVVGITIGFALKAGQIPGFRFLKVKYPRKSVNHNGFAKEMEEFPN 357
OY 352 CH-----KG-----SWCSSNOLCRECOAFMAHTMPKIKAFSMSGSA 386
Db 358 CHLOGAKGRLPVDTPLRGHESGDRFNSSTARPRPLCTGDNISSVETPIIDYTHLRIS 417
OY 387 YNAVRAVAAGHLHOLLG-----ASGACSR-GRVYPMOLBQIKHVEHLK-DTV 437
Db 418 YNVLAVYSIAHAQDITYTCLPGRGLFTNGSCADIKKVEAMQVILKHLNFTNMGEQV 477
OY 438 AFNDRPDLSSYNIIAMDMNPKMTFTVLSSTSPV-----OLNINERKIOH 486
Db 478 TFDGCGDLVGNYSIINHLSPED-----GSIVRENGVYNAVYAKKGRLEINERKILMS 531
OY 487 GKDNQVRKSVCSDDLSEGHOR-VYTGPHHCCFECVPCGAGTFLNKSJLYRCOPCGKEEMA 545
Db 532 GFSREVEFNSRDCLAGTRKGIIEGEPTECECECPDESDETDASACNKCDDPMS 591
OY 546 PEGSOTCEPRVVLALREHTSWVLLAANTLLLLLSTAGLPMHLDTPVYRAGGRLC 605
Db 592 NENTSCIAKEIEFLSWTEPFGIALTLFVAVIGIFLTAFLVFLKFRNTPIYKATNRLS 651
OY 606 FLMSGSLAAGSGSLYGFGEPTPRACILROALPLAGTIFLSCILVRSFOLIIIFKSTK 665
Db 652 YLLFSLLCGSSSLFIFGEQDMTCRLROPAFGISFLVLCISCLIVAKNRVLVE--AK 709
OY 666 VPFFYH-AMVONHGAGLFVMISSAOLILCLTWLVWTPLPAREXORPPLHMECTETN 724
Db 710 IPIFHRKRWGLNQFLVFLCTFMQIYCIWLYTAPSSRYNDELDEILFIITCHGGS 769
OY 725 --SIGFLIATLYNGLTISAFACSYLGKDPENINEAKCVTFSLFNFVSIAPFTT-AS 781
Db 770 LMAIGFLIG--YTCLLAICFFFAFKSRKLEPENNEAKFITFSMLIFFIWISFIPAYAS 827
OY 782 VYDKYLPANMMAGLSLSSGFGYFLPKCYVILCRPDLSTEHQAS 830
Db 828 TY-GREVASAEVIAIILASFGLLACIFPNKIYIILFKPSRMTIEVRCS 875

```

```

RESULT 5
US-08-353-784-7
; Sequence 7, Application US/08353784
; Patent No. 6011068
; GENERAL INFORMATION:
; APPLICANT: Edward F. Nemeth, Edward M.
; APPLICANT: Brown, Steven C. Hebert,
; APPLICANT: Bradford C. Van Wagenen, Manuel
; APPLICANT: F. Balandrin, Forrest H. Fuller,
; APPLICANT: Eric G. Delmar, and Scott T. Moe
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,784
; FILING DATE: 9 December, 1994

```

```

; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 8
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 209/069
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-353-784-7

```

```

Query Match 25.4%; Score 1142; DB 3; Length 1078;
Best Local Similarity 31.0%; Pred. No. 2e-105;
Matches 276; Conservative 163; Mismatches 358; Indels 92; Gaps 22;

```

```

OY 16 SCCW---AFACHSTRESSPD--FTLPDYLLAGLFFLHSGCL---QVRHREVTLCDRSC 66
Db 5 SCCWLLALTWHTSAYGPDQRAQKKDIIIGLFIHGVAKKDDKSRPESVBCIR-- 62
OY 67 SFNEGHYLPQAMRLGVEEINNSTALLPNTIYCYLDVYCSDSANVATLRLVSLPGOH 126
Db 63 -YNFRGFMLQAMIPAIIEINSSPALLPLNTLTGYRIFDTCNTVSK--ALEATLSFAQNK 119
OY 127 TELQ-----DLHSPVYLANIGPDSNRATTAALLSPFVPMISVAASSETISVKRQ 181
Db 120 IDSLNLDKCNSEHPISTIYAVGATGSGVSTAVANLGLFYIPOVSTVASSRLLSNKQ 179
OY 182 YPSFLRITPNKQYQETVWLLLOKFGWTWISLVGSSDDYGGOLGVALENOAGGOCICAF 241
Db 180 FKSFLRTIPNDHQATMAADITIEFRMNMVGTIADDDYGRGICIKFRFAERERICIDF 239
OY 242 KDIMPFAQVQDERMOCLMRHLAOGATVYVVFSSRQLARVEFESVLTNLTKYVASE 301
Db 240 SELI--SOYDDEEIOHVEVIONSTAKIVYFSSGPDLEPLKEIVRNITGKIMLASE 297
OY 302 AMALSRHITGVPGIORIGMVAVAIORAVPGLKAF-----EAYARADKEARPP 351
Db 298 AMASSSLIAMQYFHVVGITIGFALKAGQIPGFRFLKVKYPRKSVNHNGFAKEMEEFPN 357
OY 352 CH-----KG-----SWCSSNOLCRECOAFMAHTMPKIKAFSMSGSA 386
Db 358 CHLOGAKGRLPVDTPLRGHESGDRFNSSTARPRPLCTGDNISSVETPIIDYTHLRIS 417
OY 387 YNAVRAVAAGHLHOLLG-----ASGACSR-GRVYPMOLBQIKHVEHLK-DTV 437
Db 418 YNVLAVYSIAHAQDITYTCLPGRGLFTNGSCADIKKVEAMQVILKHLNFTNMGEQV 477

```

```

QY 438 AFNDNDPLSSYNIAMDMNGPKMTFTVLSSTWSPV-----QLINETKIQMH 486
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 478 TDECGDLVGNYSIIMWHSPEDE-----GSIVFEKVGYYNYAKKGERLFTINEEKILMS 531
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 487 GKDNOVPKSVCSDDCLEGHOR--VYGFHHCCFPCVCGAGTFLNKSDLYRCPCGGEEMA 545
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 532 GFSREVPKSNCRDCLAGRKGIIEBPTCCFCEVCPCPGEYSDETDASACNKCPCPDFFWS 591
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 546 PEGSOTCPFRVTVFLALREHTSMVLLAANTLLLLLTAGLFAWHLDPVVRASAGRLC 605
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 592 NENHNSCAKEIEFLSWTEPFGIALTLFAVLGIFLFAVLGFIKRNPIYKATNRELS 651
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 606 FLMLGSLAAGSSLYGFGEPPRPACLRQALFALGFTFLSCLYVRSQQLIIIFKFSRK 665
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 652 YLLFLSLCCFSSSLFEJLGEPODMWCRLOPAFGISFVLCISCIIVKTRVLVEE--AK 709
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 666 VPTFPH-AMVONHAGLFLFMISAAQLLCLMLVWTFPLPAREVORFPHVLWLECTEIN 724
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 710 IPTSHRKMGINLOFLVFLCTEPMOIVICVIMLTAPPSSYRNODEDELLIFITCHEGS 769
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 725 --SLGFIILAFYNGLLISAFACSYLGKDPENYNKACVTFSLLENFVSWIAFTT--AS 781
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 770 LMALGFLIG--YTCLLAICFEFFAFKSRKLPENNEAKFITFSMLIFTWISFIYAVAS 827
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 782 VYDGKYLPAANMAGISLSGFGYFLPKCYVYIICRPDLNTEHFOAS 830
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 828 TY-GREVSAYEVIAIILASFGILACIFENKIXYIILEKPSRNTIEEVRCSS 875
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 6

US-08-484-719B-7

Sequence 7, Application US/08484719B

Patent No. 6031003

GENERAL INFORMATION:

APPLICANT: Edward F. Nemech, Edward M.

APPLICANT: Brown, Steven C. Hebert,

APPLICANT: Bradford C. Van Wageningen,

APPLICANT: Manuel F. Balandrin,

APPLICANT: Forrest H. Fuller, Eric G.

APPLICANT: Delmar, Scott T. Moe

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: First Interstate World Center

STREET: Suite 4700

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS Word

SOFTWARE: FASTSEQ for Windows Version 3.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,719B

FILING DATE: 7 June, 1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993

APPLICATION NUMBER: U.S. 08/017,127

```

: FILING DATE: 12 February, 1993
: APPLICATION NUMBER: U.S. 07/934,161
: FILING DATE: 21 August, 1992
: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Douglas C. Murdock
: REGISTRATION NUMBER: 37,549
: REFERENCE/DOCKET NUMBER: 213/007
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 955-0440
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1078 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-484-719B-7

```

Query Match 25.4%; Score 1142; DB 3; Length 1078;
 Best Local Similarity 31.0%; Pred. No. 2e-105;
 Matches 276; Conservative 163; Mismatches 358; Indels 92; Gaps 22;

```

QY 16 SCCW---AFACHSTESSPD--FTLPBDYLLAGLFPPLHSGCL-----QVRRREPTVLTDRSC 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 SCCWLLALTMTWTSAYGPOAROKKGDITLGGFPIHPGVAAKQDLKSPSEVEICR-- 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 67 SFNHHGYHLFOAMRLGVEEINNSTALLPNTTLGQYDYVCSDAANYATLRLVSLPGOH 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 -YNFRFRRLQAMIFAEELINSSPALLPNTLGYRIFEDCNITRK--ALEATISFYAQNK 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 IEIQG-----DLIHSPYLVAVIGPDSTNRAATTAALSBFLVPMISYAASSFTLSVKQ 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 ISLNIDFECNCESEHIPSTIAYVAGARGSVSTAVANILGLFYIPOVSYAASSRLSNKQ 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 YPSFLRTINDKYOVETVLLQKFGMTWLSLVGSSDDYDGLGVOLLENQATGOGICIAF 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 FKSFLRTINDEHQATAMADITFERMNNWGITAADDYGRGIEKFRFEAEERDIDCF 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 KDIMPFAQVGERMQLKRLHQAAGATVYVSSRQLARVFPESVLTNLTKGVWASE 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 SELI--SQYSDEEIQHVEVIGNSTAKVIVVSSGPDLEPLKEIVRRNITGKIWLAS 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 302 AMALSHRITGVETQIRIGVNLGVAIOKRAVPGIKAF-----EAYARADKRAPRP 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 AWASSSLIAMPOYFHVVGITGFALRAGQIPGRFLPKVHPKRSVHNGFAKFEWEETEN 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 352 CH-----KG-----SWCSSNOLCRECOAFMAHTMPLKIAFMSMSA 386
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 358 CHLOEGAKGPLVDYTLRGHEESGDRFSNSSTAFLPCLGDEISSVETPIYDTILRLIS 417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 387 YNAAYRAVAVANGHLHOLLCC-----ASGACSR--GRVYPMWLLQEDIRKVFHLAK--DTV 437
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 418 YNYVLAIVSIAHNLQDITYCLPGRGLFTNGSCADIKVEAMQVLKHLRLHNFNNNGEY 477
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 438 AFNDNDPLSSYNIAMDMNGPKMTFTVLSSTWSPV-----QLINETKIQMH 486
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 478 TDECGDLVGNYSIIMWHSPEDE-----GSIVFEKVGYYNYAKKGERLFTINEEKILMS 531
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 487 GKDNOVPKSVCSDDCLEGHOR--VYGFHHCCFPCVCGAGTFLNKSDLYRCPCGGEEMA 545
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 532 GFSREVPKSNCRDCLAGRKGIIEBPTCCFCEVCPCPGEYSDETDASACNKCPCPDFFWS 591
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 546 PEGSOTCPFRVTVFLALREHTSMVLLAANTLLLLLTAGLFAWHLDPVVRASAGRLC 605
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 592 NENHNSCAKEIEFLSWTEPFGIALTLFAVLGIFLFAVLGFIKRNPIYKATNRELS 651
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

01 606 FLMGLSLAGSGSYLGFGEPTBRACILROLFALGFTIFLSCLTVSFOLIIIFKSTK 6655
02 652 YLLFLSLCCSSSLFFLGFGEQDWTCHRROPAFISIVYLCISCLIVATNRVLVEF--AK 70909
03 666 VPTFYH-AMVONHGAGLFEVMISSAOLILCTLWLVTWTPPLAREYORPPLVLMLECTETN 7244
04 725 --SGLFIILFNLGLSLISAPACSYLKGKDLPENYNEAKCYTFLSLFNFSVIAFTT-AS 7811
05 770 LMAIGFLIG--YTCLLLAICFFFAFKSKRCPENNEAKFTFMSLIFFIWISIPAYAS 8277
06 782 YVDGKYLDAANMMAGLSSLSGFCGYFLPKCYVILCPDNLNSTHFOAS 830
07 828 TY-GKFSVAVEVIALIASFELLACIFPNKIYITLFFKPSRNTIEVNCSS 875
08
09 RESULT 7
10 US-08-484-159-7
11 : Sequence 7, Application US/08484159
12 : Patent No. 6313146
13 :
14 : GENERAL INFORMATION:
15 : APPLICANT: Bradford C. Van Wagenen
16 : APPLICANT: Manuel F. Balandrin
17 : APPLICANT: Eric G. Del Mar
18 : APPLICANT: Edward F. Nemeth
19 : TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
20 : TITLE OF INVENTION: MOLECULES
21 : NUMBER OF SEQUENCES: 20
22 :
23 : CORRESPONDENCE ADDRESS:
24 : ADDRESSEE: Lyon & Lyon
25 : STREET: First Interstate World Center
26 : STREET: Suite 4700
27 : STREET: 633 West Fifth Street
28 : CITY: Los Angeles
29 : STATE: California
30 : COUNTRY: USA
31 : ZIP: 90071
32 :
33 : COMPUTER READABLE FORM:
34 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
35 : COMPUTER: IBM PC compatible
36 : OPERATING SYSTEM: PC-DOS/MS-DOS
37 : SOFTWARE: FASTSEQ
38 :
39 : CURRENT APPLICATION DATA:
40 : APPLICATION NUMBER: US/08/484,159
41 : FILING DATE: 7 June, 1995
42 : CLASSIFICATION: 435
43 :
44 : PRIOR APPLICATION DATA:
45 : PRIOR APPLICATION DATA: including application
46 : PRIOR APPLICATION DATA: described below: 9
47 : APPLICATION NUMBER: 08/353,784
48 : FILING DATE: 9 December, 1994
49 : APPLICATION NUMBER: PCT/US/94/12117
50 : FILING DATE: 21 October, 1994
51 : APPLICATION NUMBER: U.S. 08/292,827
52 : FILING DATE: 23 August, 1994
53 : APPLICATION NUMBER: U.S. 08/141,248
54 : FILING DATE: 22 October, 1993
55 : APPLICATION NUMBER: U.S. 08/009,389
56 : FILING DATE: 23 February, 1993
57 : APPLICATION NUMBER: U.S. 08/017,127
58 : FILING DATE: 12 February, 1993
59 : APPLICATION NUMBER: U.S. 07/934,161
60 : FILING DATE: 21 August, 1992
61 : APPLICATION NUMBER: U.S. 07/834,044
62 : FILING DATE: 11 February, 1992
63 : APPLICATION NUMBER: U.S. 07/749,451
64 : FILING DATE: 23 August, 1991
65 : ATTORNEY/AGENT INFORMATION:
66 : NAME: Heber, Sheldon O.
67 : REGISTRATION NUMBER: 36,119
68 : REFERENCE/DOCKET NUMBER: 214/101
69 :
70 : TELECOMMUNICATION INFORMATION:

```

```

: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1078 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: OS-08-484-159-7

Query Match      25.4%; Score 1142; DB 4; Length 1078;
Best Local Similarity 31.0%; Pred. No. 2e-105;
Matches 276; Conservative 163; Mismatches 356; Indels 92; Gaps 22.

OY 16 SCCW--AFACHSTESSPD--FTLPDGYLLAGLPFLHSGCL---QVRHREPVTLCDRSC 66
D 11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 5 SCCWLLALMTWTSYAGPDPQRAQKGDITLGLGFLHFGAAKDDKLSRPESEVICIR-- 62
OY 67 SNEGHYHLFQAMRLGVEINNSTALLRPITGLYQOLDYDCSSANVATLRSLDPCQHH 126
D 63 -NFGKPFRLQMLAIEIEINSSPLBLPLTLGYRFPDCNYSK--ALEATLSVAQAK 119
OY 127 IELQG----DLHSPYLAIVIGPDSYTRATYTAALLSPFLVPMISVAASETLSVKRO 181
D 120 IDSLWDERCNSHHIPRIAYVAGATGSGSVTAIVANLLGLFIIPOVSYASSRLLSNNKQ 179
OY 182 YPSLRITPDNDKYQVETWVLLQKFGWTWISLVGSSDDYGLQVOLAENQATGCGICIAF 241
D 180 FKSFRITPDNDHQATAMADIETFRMNVVGTIADDDYGRGPIEKFRFEAEERDIDCF 239
OY 242 KIMFSAQVGERHQCLLRHLAOGATVYVVSRSQALARVEESVLTNLGKWWAVASE 301
D 240 SELL--SQYSDEELQHYVEVQNTAKYIYVSSGPDLEPLIKETIYRNTIGKITMLASE 297
OY 302 AWALSRHITGVYGIQIGWLGVAIQKRAVGLKAF-----EAYARADKEDAPRP 351
D 298 AWASSLIAMPOYFHVVGTTIGFALKAGIPEGRFEFLKVHPRKSVHNHFAKEFWEETFN 357
OY 352 CH-----G-----SNCSNOLCECOAFNAHMPKAKASMSA 386
D 358 HLLQRCAGAPLPVDYTFPLKHEHSGDRFSSNSTAFRRPLCTGDDNISSEVETPYDIYHLRIS 417
OY 387 YNAAYAAVAHAGHLQQLGCG-----ASGACSR--GRVYPMQLDQIHKVFHLRK--DTV 437
D 418 YVVLVAYIASIALADQITCLPGRGLFTNGSCADIKKYEAQVULKHLRLHRLNTFTNMGEQY 477
OY 438 AFENKRDPLSSNITIAMDMNGPKWTFYVLGSSTWSPV-----QLINTEKIQWH 486
D 478 TFEDEGDIAGVNSIINMHLSPD-----GSITVEFEGYVNYAKKGERLFTNEKILMS 531
OY 487 GKDNQVPKSVCSDDLEGHOR--VVTGFHHCCFECYPCAGFELNSDLYRCQPCGKEEWA 545
D 532 GSRREVPSNCRDCLAGIRKGIIGSEPRCCGCEYECRPGESDSDTDSACKKCPDDEFS 591
OY 546 PEGSOTCEPRVYVTLALREHITSWLLAANTLLLLILGLAGLPANHLDTPVYRSAGRLC 605
D 592 NENHNSCKIAKELFESWTEPFGIATLPAVLGIFLTAFLVGLFIFKFRMPRIKVAKNRELS 651
OY 606 FLMLGSLAAGSGSLGFPEEPTRPAPCLRQALFALGFTFLSCFLVRSQOLIIFKFSFK 665
D 652 YLLLSLCCRESSSLFTEIGEPDMTCRLRQAPFGISFVLICISLIYKTRVLVE--AK 709
OY 666 VTFYH-AMVONHAGGLFVMISSAAQLLCLTALVWVTPPLPAREYQRPBHYMLECTETN 724
D 710 IPTSHRKWGMGINLQFLVFLCTFMQIYICVIMLTAPASSYRNOLEDEEIFICHESG 769
OY 725 --SLGPIALFLNGLLSISAFASVYLGKDPENYMEAKCVTSLLFPNFSIAFTT--AS 781
D 770 IMALEPLLG--YTCLLAACEFFAKSRKLPENEDEAKTISMLIFLFWLSFTPAYAS 827
OY 782 YVDGYTLPAANMAGLSLSSGFGYVFLPKCVIILCRPLNTEHFOAS 830

```

Db 828 TY-GKFSVAVEVIAIILAAISGLIACIFENKRIYIILFKPSRNTIEEVCS 875

RESULT 8

US-08-485-588-6
Sequence 6, Application US/08485588

Patent No. 5688938

GENERAL INFORMATION:

APPLICANT: Edward M. Brown

APPLICANT: Steven C. Hebert

APPLICANT: Forrest H. Fuller

APPLICANT: James E. Garrett, Jr.

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: First Interstate World Center

STREET: Suite 4700

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,588

FILING DATE: 7 June, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: Including application

PRIOR APPLICATION DATA: described below: 9

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993

APPLICATION NUMBER: U.S. 08/017,127

FILING DATE: 12 February, 1993

APPLICATION NUMBER: U.S. 07/934,161

FILING DATE: 21 August, 1992

APPLICATION NUMBER: U.S. 07/834,044

FILING DATE: 11 February, 1992

APPLICATION NUMBER: U.S. 07/749,451

FILING DATE: 23 August, 1991

ATTORNEY/AGENT INFORMATION:

NAME: Heber, Sheldon O.

REGISTRATION NUMBER: 38,179

REFERENCE/DOCKET NUMBER: 213/005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 1088 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-485-588-6

Query Match 25.2%, Score 1130, DB 1, Length 1088,

Best Local Similarity 30.8%; Pred. No. 3.3e-104;
Matches 277; Conservative 162; Mismatches 358; Indels 102; Gaps 23;

QY 16 SCCW---AFACHSESSPD--FTLPDYLALGFPPLHSGL-----QVRHREPTLDCRSC 66
Db 5 SCCWLLALTWHTSANGPQRAKKGDIILGIFPIHFGVAAKDQIKSRPSEVCEIR-- 62
QY 67 SFNEHGYHLFOAMRLGVEEINNSTALLPNTTGQYQDYVDCSDSANVATLRYLSLGGQH 126
Db 63 -YNEFGFRLQAMIFAEIEINSSPALLNLTGRIEDPCNMYSK--ALEATLSFYAQNK 119
QY 127 IELQG-----DLHSPYVLAVIGPSTNRATALLSPFLVPMIYSTAASSFTLSVKQ 181
Db 120 IDSLNDEFQNCSEHLPSTLAVGATGVSVAANMLGLFYIPQYVASSSLSSLNKKQ 179
QY 182 YPSFLRTIPNDKYQVETWVLLQKFGWYSLVSGSDPDYQGLQALQENQOGICIAF 241
Db 180 FKSFLRTINDEHQATMADIIFFRMNVGTIAADDYGRGIEFREFAEERDQICDF 239
QY 242 KDIMPFAQVGERMQLKRLHQAQATVYVFFSSQRLARVFESVYLNTLGRVAVASE 301
Db 240 SELI--SOYSDEEIQHVEVQIONSTAKVIYVSSGPDLEPLIKETIVRRNTGKIWLASE 297
QY 302 AMALSHHTIGVCIQRIQAVLQRAVPGIKAF-----EAYARADKRAPRP 351
Db 298 AMASSSLIAMPOYFHVYGGTIGFALKAGQIPGREFELKRVHPKSVHNGFAKEEWEETEN 357
QY 352 CH-----KG-----SMCSNOLCREQAFMAHMPKLFKAFASSSA 386
Db 358 CHQEGAKGPLVDYTLRGHEESGDRFSNSTAFRLCTGDEMISSVEPIYDITLRLRS 417
QY 387 YNAAYRAVAVAHGLHQLLCC-----ASGACSR-GRVYPMQLLEQIHKVHFLLR-DFV 437
Db 418 YNYVLAVYSIAHLADQIYICLPGRGLFTNGSCADIKKVAQVILKRLHLENNMGEOY 477
QY 438 AFQNDNDPLSSYNTIAMDNNGRKMTFTVIGSSWSVY-----QNTINETIQNH 486
Db 478 TDECDGLVGNYSIIMWHLSPED-----GSIVFEKVGYYNYAKKGERLFIKEEKILMS 531
QY 487 GKDN-----QVPSKVSQSDCLEGHQR-VVTGFHHCCEFCVCGAFTLNSKDLXR 535
Db 532 GFSREPLTVLSVQVFPSCNSRDCLAGTRKGIIEBPCCFCFCVCPDGEYDETDASA 591
QY 536 COPCGKEAPREGSQTCPRTYVFLALREHTSVLLAANTLLLLLGAGLFAWHLDPP 595
Db 592 CNKCPDPEFMSNENHTSCIAKEIFLSWTEPGIALFLFVAGLFLVAFVIGYIKFRNRP 651
QY 596 VYRSAGRICFLMLGSLAGSGSLYGFGEPTRPACLLQALFALGFTIFLSLTYSRQ 655
Db 652 YKATRELSYLLFLSLCCFSSSLFEIGEPDWTRLROPAPGISFVLCISQILVKTNR 711
QY 656 LIIIFKSTKVPFEH-AWVONHAGLFWMISSAOLILCLWLVWTPLPAREYORFPH 714
Db 712 VLVFE--AKIPFSHRKRWGLNLOFLVFLCTPMQIVLCVILYAPSSSYNQLDEDE 769
QY 715 LVMLECTETN--SLGFLIALFLYNGLLISAFACSYGLKDPENYNAKCVTESLNFVS 772
Db 770 IIFITCHEGSLMALGFLIG--YTCLLAICFFFAFSRKLPEFNFAKPLTFSMLLEFY 827
QY 773 WIAFFTT-ASVIDGKYLPAANMAGLSLSSGGGFLFKCYIILCRPLDNSTENHQAS 830
Db 828 WISFIPAVASTY-GKFSVAVEVIAIILAAISGLIACIFENKRIYIILFKPSRNTIEEVCS 885

RESULT 9

US-08-484-565-6

Sequence 6, Application US/08484565

Patent No. 5763569

GENERAL INFORMATION:

APPLICANT: Edward M. Brown

APPLICANT: Steven C. Hebert

APPLICANT: James E. Garrett, Jr.

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE


```

? TITLE OF INVENTION: MOLECULES
? NUMBER OF SEQUENCES: 20
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Lyon & Lyon
? STREET: First Interstate World Center
? STREET: Suite 4700
? STREET: 633 West Fifth Street
? CITY: Los Angeles
? STATE: California
? COUNTRY: USA
? ZIP: 90071
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: FASTSEQ
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/484,565
? FILING DATE: 7 June, 1995
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? PRIOR APPLICATION DATA: including application
? PRIOR APPLICATION DATA: described below: 9
? APPLICATION NUMBER: 08/353,784
? FILING DATE: 9 December, 1994
? APPLICATION NUMBER: PCT/US/94/12117
? FILING DATE: 21 October, 1994
? APPLICATION NUMBER: U.S. 08/292,827
? FILING DATE: 23 August, 1994
? APPLICATION NUMBER: U.S. 08/141,248
? FILING DATE: 22 October, 1993
? APPLICATION NUMBER: U.S. 08/009,389
? FILING DATE: 23 February, 1993
? APPLICATION NUMBER: U.S. 08/017,127
? FILING DATE: 12 February, 1993
? APPLICATION NUMBER: U.S. 07/934,161
? FILING DATE: 21 August, 1992
? APPLICATION NUMBER: U.S. 07/834,044
? FILING DATE: 11 February, 1992
? APPLICATION NUMBER: U.S. 07/749,451
? FILING DATE: 23 August, 1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Heber, Sheldon O.
? REGISTRATION NUMBER: 38,179
? REFERENCE/DOCKET NUMBER: 213/006
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (213) 489-1600
? TELEFAX: (213) 955-0440
? TELEX: 67-3510
? INFORMATION FOR SEQ ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1088 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-484-565-6

```

Query Match 25.2%; Score 1130; DB 1; Length 1088;
 Best Local Similarity 30.8%; Pred. No. 3.3e-104;
 Matches 277; Conservative 162; Mismatches 358; Indels 102; Gaps 23;

```

QY 16 SCCW---AFACHSTRESSPD--FLTPGDLVLAQFPLSGCL---QVYRHREYVTLCDRSC 66
DB 5 SCCWLLATLWHTSAVPPDRAOKKGIILGGEPHFHGAAKQODLKSRESVECTR-- 62
QY 67 SFNEGHVLFQAMRIGVEYEINNSTALPNTLTGQYQDVSDSANVYATLRLSLPCQH 126
DB 63 -YNRGRFWMQAMIFALEEINSSPALLPNTLTGRIEDTCNTVSK--ALENTLSFVQNK 119
QY 127 IEIAG-----DLHYSTPTVLAIVGPDSTNRAATTAALLSPFLVPMISTYASSELVSKRQ 181
DB 120 IDSINIDFPCNCEHIPSTIAVVGATGSGVSTAVANLIGLFYIPVSVASSRLLSNKQ 179

```

```

QY 182 YPSFLRTIPNDKXQVETWVLLQKRGWTWISLVGSSDDYDQGLQALBNQATGQICIAF 241
DB 180 FKSFLRTIPNDHQATADITETFRMNMVGTITAADDYGRGILEKFEEDBERDICTIDF 239
QY 242 KDIMPFSQVODERMOCIMRLHAOGATVAVVVFSSRQLARFFESVVTJNLTKWVASE 301
DB 240 SELT--SQYSDDEEIQHVEVIONSTAKIVYVFSGPDLEPLIKEIVARNITGKMLASE 297
QY 302 AMALSRHITGVPGIQRIGMVLGVAIQKRAVPGIAF-----EBAVARADKEADRP 351
DB 298 AMASSSLAMPQYFHVVGITGFALKAQOIPGFRFLKAVPRKSVHNGFAKEFEETFN 357
QY 352 CH-----KG-----SMCSSQOLCEQOAFMAHMPILKAFMSGA 386
DB 358 CHLOGBAKPLVPDTFLRGHEESGDRSNSSTARPPDLCTGDNISSVTPYIDTDLAIS 417
QY 387 YNAYRAVYAVAHGLHQLGQ-----ASGACSR-GRVYPMQLLEQIHKVFLHKK-DTV 437
DB 418 YNVLAVYSIAHALQDIYTCLPGRGLFTNGSCADIKKVEAMQVILKHLHMLPTNMGEQV 477
QY 438 AFNDNRDPLSSYNTIAMDNGPKMTFTVYLGSTSPV-----QININETKIOMH 486
DB 478 TFDECGDLVGNYSIIMHLSPEDE-----GSIVREKGVYVYAKKGRRLFINEKILMS 531
QY 487 GKDN-----QVPRKSVCSQCLEGHOR-VYTGFFHCCQECVPCGAGTFUNKSDLYR 535
DB 532 GFSEPEPLTVLSVLQVPPSNCSRDCLACTRKGILIEGPTCCFEVECDGEDYSTDTDSA 591
QY 536 COPCKEEMAPESQOTCPFRIVFLALREHTSWVLLANTLLLLTAGLFAWHLDP 595
DB 592 CNKCPDDEWSNMHTSCIAKEIEFLSWTEPFGIALTLFAVIGIFLTAFLVGFIFRMT 651
QY 596 VVRBAGRGFLCMGLSAGSGSLXGFGEPTRRACLRLQALFALGFIIFLSCLTVRSFQ 655
DB 652 IVKATNRLSYLLFSLCCPSSSIFLGEPODWTCLRQAPGISFVLCTISCIIVKTR 711
QY 656 LIIFKFEKTYPTFYH-AMVONHGAGLFVMISSAQAOLICTLTVMPPLAREQRP 714
DB 712 VLVFPE-AKIFPSHRKMWGLNLOFLVFLCTFMQIYICVWLTAPSSRYNDELDE 769
QY 715 LVMLECTETN--SLGFIAPLYNGLISAPACSYLGKDLDPENYNEAKCVFSLLENVS 772
DB 770 IIFTCHEGSLMALGFLIG--YTCLLAICFFFAKRSKRLDENNEAKFITFSMLIF 827
QY 773 WIAFFT-ASYVDGKIYLPANMMAGLSLSSGFGYFLPKCYVILCRDLNSTEHPQS 830
DB 828 WISFIPAYASTY-GKFSVAVEVIALAASFGLACIFPNKIYIILFKPSRMTIEEVRS 885

```

RESULT 10
 US-08-480-751-6
 Sequence 6, Application US/08480751
 Patent No. 5858684
 GENERAL INFORMATION:
 APPLICANT: Edward F. Nemeth
 APPLICANT: Edward M. Brown
 APPLICANT: Steven C. Hebert
 APPLICANT: Forrest H. Fuller
 APPLICANT: James E. Garrett, Jr.
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 TITLE OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:


```

1      FILING DATE: 21 October, 1994
2      APPLICATION NUMBER: U.S. 08/292,827
3      FILING DATE: 23 August, 1994
4      APPLICATION NUMBER: U.S. 08/141,248
5      FILING DATE: 22 October, 1993
6      APPLICATION NUMBER: U.S. 08/009,389
7      FILING DATE: 23 February, 1993
8      APPLICATION NUMBER: U.S. 08/017,127
9      FILING DATE: 12 February, 1993
10     APPLICATION NUMBER: U.S. 07/934,161
11     FILING DATE: 21 August, 1992
12     APPLICATION NUMBER: U.S. 07/894,044
13     FILING DATE: 11 February, 1992
14     APPLICATION NUMBER: U.S. 07/749,451
15     FILING DATE: 23 August, 1991
16     ATTORNEY/AGENT INFORMATION:
17     NAME: Heber, Sheldon O.
18     REGISTRATION NUMBER: 38,179
19     REFERENCE/DOCKET NUMBER: 213/006
20     TELECOMMUNICATION INFORMATION:
21     TELEPHONE: (213) 489-1600
22     TELEFAX: (213) 955-0440
23     TELEX: 67-3510
24     INFORMATION FOR SEQ ID NO: 6:
25     SEQUENCE CHARACTERISTICS:
26     LENGTH: 1088 amino acids
27     TYPE: amino acid
28     TOPOLOGY: linear
29     MOLECULE TYPE: protein
30     US-08-943-986-6

```

```

0y 16 SCW-----AFACHSPRESSD--FTLPQVYLAGLEPRLSGCL-----OVRNHPEVTLCDRSC 66
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 5 SCWVLLALVHTHTSYGPDQAOAKKQDIIAGGLEPPIHFVGAAKODUKSRESPECTR-- 62
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

0y 67 SFNEHGYYLFOAMRLGVEEINNSTALLPNTLYQLQYDVCSDSANVYATLRVLSLPGOH 126
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 63 -YNFRGFWMLOAMFATEEINSSPALLPNTLYGRIFDTQWYK--ALEATISFVAQNK 119
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

0y 127 TELG-----DLHSPVYLAVIDPDSNTRRATYTAALLSPFLYPMISYAASSETLSYKRO 181
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 120 IDSJLNLDEFNCSEHISTPSTIAVWGATSGSVTAANLLGLFIYIPQVYASASSRLSNKQ 179
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

0y 182 YPSFLRTIPDNKQVETPMVLLQKFKGWTYSLVSSDDQGLQVQALENATGGCTIAF 244
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 180 FKSPFLRTIPDNHQTAAADIIIEFFRNMWGTIAADDYGGPFGTEKFEEREAEEDCIDF 239
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

0y 242 KDIMEFSQOVDERMOCMLRHQAQATVVVYFSSROLARFEFSVLYTNJTGKVVWASE 301
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 240 SELI--SOYSDEELQHVVEYTONSTAKVIVYESSGPDLEPLKEIYRNTITKIMLASE 297
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

0y 302 AMALSRIHTGVPQIORIGMVJGVAIQRKRAVPGCLAF-----BEAYARADKEAPR 351
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 298 AMASSLIAMPQYFHVVGCTIGFALKAGQIHPREFPLKVNPRKSVANGFAKEWMETFN 357
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

0y 352 CH-----KG-----SWCSSNQLCRECOAFMAHTPKLKAESSSA 386
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 358 CHLOEGAGPLPVDTFLRGHEESDGRFSNSSTARPLCTGDENISSEVETPYIDYTHLRIS 417
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

0y 387 YNAVRAYVAVHAGHJOLLGC-----ASGQCSR-GGVYPMQOLLEQJHXKHVFLHK-DYV 437
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 418 YNVLYAVYSIHAALIDYITCLPGRGLEFTNGSCADIKKVEAMOVYKLHRLHFNFTNNKEOY 477
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

0y 438 AFNDNRPLUSSYNIILADMDNPKWTFTVLGSSWTSPV-----OLNINETKIOWH 486
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 478 TFEDEGDLVGNYSIINHLSPED-----GSIVKEVGYIYNVYAKKBERFLFINDERKLMS 531
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

0y 487 GKDN-----QVPRKSVCSDDCLEGHOR-VVTGFHHNCCFECVPCGAGTFLNKSOLDYR 535
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

```

12
US-08-353-784-6
Sequence 6, Application US/08353784
Patent No. 6011068
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth, Edward M.
APPLICANT: Brown, Steven C. Hebert,
APPLICANT: Bradford C. Van Wageningen, Manuel
APPLICANT: F. Balandrin, Forrest H. Fuller,
APPLICANT: Eric G. DelMar, and Scott T. Moe
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,784
FILING DATE: 9 December, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 8
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451


```

QY 67 SENEHGYHLEQAMRLGVEINNSTALPNTITLGYQLYDVCSDSANVATLRLVSLPGQH 126
DB 63 -YNFGRFRLQAMTFAIEBINNSPALLPRLTLGRIEDCNVSK--ALEATLSFAQMK 119
QY 127 IELQ-----DLHYSPTYLAVIGPDSINRAATTAALLSPFLVPMISVASSSETLSVKRO 181
DB 120 IDSLMDEFCNCEHPISTIAVVGATGSGVSTAVANLLGLFYIPQVYASSSRLSLSNKNO 179
QY 182 YPSFLRTIPNDKYQVETWLLLOKFGMTWISLVGSSDDYGOLGQVQALEMQATGOGICAF 241
DB 180 FESFLRTIPNDHQATAMADITIEFRMNVGTIAADDDGRPGIEKFRREAEERDICTDP 239
QY 242 KDIMPESAQVGERMOCMLRHLAQAGATVYVSSRQLARVEFESVLTNLTKVWASE 301
DB 240 SELI--SQYSDDEEIQHVEYVIONSTAKYIVFSSGPDLEPLIKEIVRNITGKILASE 297
QY 302 AMALSRHTGVGIGRIGVLAIOKRAVPLKAF-----EEAYARADKAPRP 351
DB 298 AMASSLLAMPQYFHVVGGTIGFALKAGQIPGRFELKRVHBRKSVHNGFAFEWEETEN 357
QY 352 CH-----KG-----SMCSNOLCRECOAFMAHTMPKLFKAFMSA 386
DB 358 CHLOGAGKPLVDYFRLGRHESSGDRFSNSTAFRLCTGDNISSEVETPYIDYTLKRS 417
QY 387 YNAYAVYAVAHGLHQLLGC-----ASGACSR-GRVYPMOLLEQIKRVHFLHK-DTV 437
DB 418 YNVLAVYVIAHALQDIYCLPGRGLFTNGSCADIKKIVAMQVILKRLHNTFNNGEQV 477
QY 438 AFENDRDPSSNYIIAMDMNGPKMTFTVLSSTSPV-----OLNINETKIOMH 486
DB 478 TFDECGDLVGNSTINMHLSPED-----GSTVFEVGYVNYAKKGERLFTNEEKILMS 531
QY 487 GKN-----QVPKSCSSDCELEHOR-VVTGFHHCCECVPCGAGTFLNKSIDLX 535
DB 532 GFSRPLFVLVSLQVFPNSCRDLACGRKGIIEBEPCECEVCBPGGEVSDERDASA 591
QY 536 COPCKEMAPRGSOCTCPRTVVFALREHTSMVLIANTLTLTLGLTAGLFPAMHLDTP 595
DB 592 CNKCDDEFSNENHSCIAKEIEFLSWTEPFGLALFLFAVLGIFLAFVIGFIRKPNP 651
QY 596 VYRSAGRCLFELMGLSIAAGSIVGFGEPRPACLLRQALFALGFTIFLSCLTVRSQ 655
DB 652 YKATNRELSTLRLSLCCFSSSLFIEPODMTCRLQPAFGISFVLCISLIVKTRN 711
QY 656 LIIRKFTKVPFYH-AWVQNHGAGLFYMISSAQLICTLWLVWTFPLPAREYORFPH 714
DB 712 VLVFE--AKIPTSRKMWGLNQLFLVFLCTPMQIVICVIMLVITAPSSVRNQLDEDE 769
QY 715 LVMECTEN--SLGFIATFLYNGILSLISAFACSTLGDLPENYNEAKCVTSLSLFNYS 772
DB 770 IIFICHGSLMALFLIG--YTCLIAACFPFAKSRKLPENFNFAKIFTSMLFFTY 827
QY 773 WIAFTT-ASYVDKYLLPANNMAGLSSISGFGVFLPKCVIICRPLNSTEHQAS 830
DB 828 WISFLPASTY-GKFSVAVEVIALIASFGLLACIFENKIYIILFKPSRNTIEVRS 885

```

RESULT 15

US-09-134-513-2

Sequence 2, Application US/09134513

Patent No. 6210964

GENERAL INFORMATION:

APPLICANT: Brown, Edward M.

APPLICANT: Diaz, Ruben

APPLICANT: Bai, Mei

APPLICANT: Quinn, Stephen J.

TITLE OF INVENTION: The Avian Extracellular Calcium-Sensing

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESS: Vinson & Elkins L.L.P.

STREET: 1455 Pennsylvania Avenue, N.W.

```

CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,513
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: BR1331/13003
TELEPHONE: (202)639-6604
TELEFAX: (202)639-6604
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1059 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-09-134-513-2

```

Query Match 25.1%; Score 1127.5; DB 4; Length 1059;
Best Local Similarity 30.7%; Pred. No. 5,6e-104;
Matches 276; Conservative 163; Mismatches 352; Indels 107; Gaps 24;

```

QY 14 LISC-----WAFCHTSSPDITLPGDYIALFLHSCSCL-----QVRHREPTVC 62
DB 3 LYSCEILLFTWNTAAYGPNQRAQ--KKGDITLGLGFLPHGVAAKDDLSRPSVBC 60
QY 63 DSCSSENEHGYHLEQAMRLGVEINNSTALPNTITLGYQLYDVCSDSANVATLRLVSLP 122
DB 61 IR---YNFGRFRLQAMTFAIEBINNSPALLPRLTLGRIEDCNVSK--ALEATLSV 115
QY 123 GOHHEILOG-----DLHYSPTYLAVIGPDSINRAATTAALLSPFLVPMISVASSSETLS 177
DB 116 AONKIDSLMDEFCNCEHPISTIAVVGATGSGVSTAVANLLGLFYIPQVYASSSRLSL 175
QY 178 VKROYSPFLRTIPNDKYQVETWLLLOKFGMTWISLVGSSDDYGOLGQVQALEMQATGOGI 237
DB 176 NKNQFSPFLRTIPNDHQATAMADITIEFRMNVGTIAADDDGRPGIEKFRREAEERDI 235
QY 238 CIAFKDIMPESAQVGERMOCMLRHLAQAGATVYVSSRQLARVEFESVLTNLTKVW 297
DB 236 CIDFSELI--SQYSDDEEIQHVEYVIONSTAKYIVFSSGPDLEPLIKEIVRNITGKIL 293
QY 298 VASEANALSRHTGVGIGRIGVLAIOKRAVPLKAF-----EEAYARADK 347
DB 294 LASEMASSLLAMPQYFHVVGGTIGFALKAGQIPGRFELKRVHBRKSVHNGFAFEWE 353
QY 348 APRPCH-----KGSWSSN-----OLCRECOAFMAHTMPKLFKAFMSA 383
DB 354 EFNFCYLPSESNNSPASAFPHRAHEBGLAGNGTAARPPCTGDNITSEVETPYMDFTL 413
QY 384 SSAYNAYRAVYAVAHGLHQLLGC-----ASGACSR-GRVYPMOLLEQIKRVHFLHK- 434
DB 414 RISTYNYLAVYVIAHALQDIYTCPECKGLFTNGSCADIKKIVAMQVILKRLHNTFN 473
QY 435 DIVAFNDNDPSSNYIIAMDMNGPKMTFTVLSSTSPV-----OLNINETKI 483
DB 474 EYVDFDEFGDLVGNSTINMHLSPED-----GSVVEEVGYVNYAKKGERLFIENKI 527
QY 484 QNHGKDNQPKVSCSSDCELEHOR-VVTGFHHCCECVPCGAGTFLNKSIDLRCOPCGE 542

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 09:14:08 ; Search time 61.51 Seconds

(without alignments)
1518.665 Million cell updates/sec

Title: US-09-819-946-2

Perfect score: 4493

Sequence: 1 MLCIARLVGLQLISCCWA.....NTEHPQASIDYTRRCGST 841

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4493	100.0	841	22	AAE11969 Human novel G-prot
2	4489	99.9	841	22	AAE10372 Human taste recept
3	4068.5	90.6	777	21	AAV45023 Human sensory tran
4	4058	90.3	763	22	AAE11970 Human novel G-prot
5	3361.5	74.8	840	21	AAV45021 Rat sensory transd
6	3359.5	74.8	840	21	AAV45026 Rat sensory transd
7	3358.5	74.7	840	21	AAV45027 Rat sensory transd
8	3357.5	74.7	840	21	AAV45028 Rat sensory transd
9	3348.5	74.5	842	21	AAV45022 Mouse sensory tran
10	1803	40.1	366	22	AAE11971 Human novel G-prot
11	1588	35.3	320	22	AAW25634 Human protein sequ

12	1370	30.5	843	21	AAV77556 Rat GPCR-B4 polype
13	1350	30.0	843	21	AAV77557 Mouse GPCR-B4 poly
14	1238	27.6	234	22	AAE11972 Human novel G-prot
15	1170	26.0	858	22	AAE10371 Rat taste receptor
16	1155	25.7	852	22	AAU08996 Human G protein-co
17	1142	25.4	1078	17	AAW11889 Human parathyroid calciu
18	1142	25.4	1078	19	AAW54846 Human parathyroid
19	1142	25.4	1078	19	AAW38274 Human parathyroid
20	1142	25.4	1078	20	AAV28840 Human calcium rece
21	1142	25.4	1078	20	AAV41780 Human parathyroid
22	1142	25.4	1078	20	AAW89565 Human parathyroid
23	1142	25.4	1078	21	AAV51827 Human calcium rece
24	1142	25.4	1078	21	AAV70325 Human wild type ca
25	1142	25.4	1078	22	AAV4391 Protein encoded by
26	1133	25.2	850	22	AAE10366 Human taste-cell-s
27	1130.5	25.2	1078	22	AAU02195 Cynomolgus monkey
28	1130	25.2	1088	17	AAW11888 Human parathyroid calciu
29	1130	25.2	1088	19	AAW54845 Human parathyroid
30	1130	25.2	1088	20	AAV41779 Human parathyroid
31	1130	25.2	1088	20	AAW89564 Human parathyroid
32	1130	25.2	1088	21	AAV51826 Human calcium rece
33	1130	25.2	1088	22	AAV74390 Protein encoded by
34	1129	25.1	1088	19	AAW38273 Human parathyroid
35	1127.5	25.1	1059	22	AAU00508 Chicken calcium-se
36	1126.5	25.1	1085	19	AAW54844 Bovine parathyroid
37	1126.5	25.1	1085	19	AAW38272 Bovine parathyroid
38	1126.5	25.1	1085	20	AAV41778 Bovine parathyroid
39	1126.5	25.1	1085	20	AAW89563 Bovine parathyroid
40	1126.5	25.1	1085	21	AAV51825 Bovine calcium rec
41	1123	25.0	1079	19	AAW54847 Rat kidney calcium
42	1123	25.0	1079	19	AAW38275 Rat kidney cell ca
43	1123	25.0	1079	20	AAV41781 Rat parathyroid ca
44	1123	25.0	1079	20	AAW94928 Rat kidney extrac
45	1123	25.0	1079	20	AAW89566 Rat parathyroid ca

ALIGNMENTS

RESULT	1
AAE11969	AAE11969 standard; Protein: 841 AA.
ID	AAE11969;
AC	AAE11969;
XX	
XX	18-DEC-2001 (first entry)
DE	Human novel G-protein coupled receptor (NGPCR) protein #1.
DE	Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW	antitense-therapy; signal transduction; behavioural disorder; obesity;
KW	heartbeat rate; inflammation; immune disorder; diabetes; cancer;
KW	coronary disease.
XX	
XX	Homo sapiens.
OS	
XX	
XX	Key location/Qualifiers
FT	Misc-difference 107 /note= "Encoded by TYT"
FT	Misc-difference 372 /note= "Encoded by RCA"
FT	
XX	
PN	WO200172842-A2.
XX	
PD	04-OCT-2001.
XX	
PF	28-MAR-2001; 2001WO-US09996.
XX	
PR	28-MAR-2000; 2000US-192978P.
XX	
PA	(LEXI-) LEXICON GENETICS INC.
XX	
PI	Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;

PI Zambrowicz B, Sands AT;
 XX MPI: 2001-616474/71.
 DR N-PSDB: AAD19501.
 XX
 PT Novel isolated polynucleotides encoding human G protein coupled
 receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 expressed NGPCRs for diagnosis of disease, and as probes or primers -
 XX
 PS
 Claim 3: Page 72-74; 80pp; English.

CC The present sequence is human novel G-protein coupled receptor
 CC (NGPCR) protein. NGPCRs are transmembrane proteins that span
 CC the cellular membrane and are involved in signal transduction after
 CC ligand binding. The NGPCR polynucleotide sequences, are useful
 CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heartbeats rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.
 CC
 XX
 XX Sequence 841 AA:

Query Match 100.0%; Score 4493; DB 22; Length 841;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLCTARLVGLDLISCCMAFACHSTESSPDTLPDGYLLAGLPLHSGCLQVRRPEVT 60
 DB 1 mlctarlvglldliscmwfachstesspdtlpgdyllaglfphsgclqyrhpervt 60
 QY 61 LCRSGSPNHHGHVLFQARMLGVEETINNTALLPNTIGQGLDVCSDSNVATLRVLS 120
 DB 61 lcrsgspnhhghvlfqarmlgveeinntallpntlgyqldvcsdsanvyltrvls 120
 QY 121 LFGQHIEIQGLDLAHSPYTLAVIGPDSINRAATTAALLSPFVPMISYASSETLSVKR 180
 DB 121 lfgqhieiqgdlahspvylavigpdsinraatallspflvpmisyaasetlsvkr 180
 QY 181 QYPSLRTIPNDKYQVEYEWVLLQKFGWTWISLVGSSDDYGLGVQALENQATGGGICIA 240
 DB 181 qypslrtipndkyoveyewvllqkfgwtwislvgssddyglgvqalenoatgggicia 240
 QY 241 FKQIMFSAOVGDERMOCIMRHIAAGATVYVVFSSROLARFEESVLTNLTKGWVMS 300
 DB 241 fkdimpfsavgdermqcmrthlaqagatvvvfssrglarvfesvltnltkgwvms 300
 QY 301 EAWALSRRHTTGVGIORIGVILGVAIQKRAVPGLKAFEEYARADREAPRCHKSGWCSS 360
 DB 301 eawalsrhttgvgiorigviltgvaioqkravpglkafeeayaradreaprchksgwc 360
 QY 361 NOLCRECOAFMAHTPRKLAFSMSAYNAVRAVAVAHGILHQLLGASGASGRVYPMW 420
 DB 361 nqlcrecoafmahmpkklafmsasaynavrayavahghlqllgcasgascgrvympw 420
 QY 421 LLEQHKVFFLHKDVTVAFNDRDPLSSYNIITAMWNGPKWTFVYVYSSYVQNLIN 480
 DB 421 lleqhkvvflhkdvtafnndrplssyniitawngpkwtvtyvysvynlind 480
 QY 481 TKIOWHGKNOVYKSVCSDDLGHORVYVGFHHCCECPGAGTFLNKSDDLYRCOPG 540
 DB 481 tkiowhgknoykvsvcsddclghorvvtgfhccfecvpgagtfllnksdlyrcopg 540

QY 541 KEEMAPESGQCEPRTVVELALREHTSWLLAANTLLLLLLGTAGLFAWMLDTPVYRSA 600
 DB 541 keemapesgqctprtvelalrehtswllaanlllllllgtaglfawhldtpvyrsa 600
 QY 601 GGRLCFLMGLSLAAGSGSLXGFFGEPTRRACLLRQALFALGFTIFLSCITVASFQIITF 660
 DB 601 ggrlcfmlgslaaagslxygfgeptrracllrqalfalgtiflscitvssfqiitf 660
 QY 661 KFSTKVPTFYHAWVONHGALEFYMISAAQLLCLTLWLVWMPPLPAREYQRPPLVMEBC 720
 DB 661 kfstkvpftyhawnhgaelfymisaaqlcltlwlvwmpplpareyqrfplvmelec 720
 QY 721 TETNSLGFILAFLYNGLSLISAFACSYLGKDLPEYNEAKCVTFSLFNFVSWIAFETTA 780
 DB 721 tetnslgfilafllynslisafacsylgkdlpeyneaekvtfslfnfvswiafetta 780
 QY 781 SVYDGKYLPAANMAGLSLSSFGGYFLPKCYVILCRDNLNSTERFOASIDYTRRCS 840
 DB 781 svydgkylpaanmaglsissfggyflpkcyvilcrpdlntstehfqsldytrrcgs 840
 QY 841 T 841
 DB 841 t 841

RESULT 2
 AAE10372
 ID AAE10372 standard; Protein; 841 AA.
 XX
 AC AAE10372;
 XX
 DT 10-DEC-2001 (first entry)
 XX
 DE Human taste receptor, hTIR1 protein.
 XX
 KW Human; taste-cell-specific G protein-coupled receptor; hTIR1; drug;
 KW genetic modulation; pharmaceutical; taste sensation; food industry;
 KW chemosensory transduction.
 XX
 OS Homo sapiens.
 XX

Key location/Qualifiers
 FT Misc-difference 274 /note="Encoded by TTTT"
 FT
 PN WO200166563-A2.
 XX
 PD 13-SEP-2001.
 XX

XX 07-MAR-2001; 2001WO-US07265.
 XX
 XX 07-MAR-2000; 2000US-0187546.
 XX
 XX 07-APR-2000; 2000US-0195536.
 XX
 XX 06-JUN-2000; 2000US-0209840.
 XX
 XX 23-JUN-2000; 2000US-0214213.
 XX
 XX 17-AUG-2000; 2000US-0226448.
 XX
 XX 03-JAN-2001; 2001US-0259227.
 XX

XX (SENO-) SENOMYX INC.
 XX
 XX Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
 XX
 XX MPI: 2001-582267/65.
 XX
 XX N-PSDB: AAD17515, AAD17516.
 XX

PT New mammalian taste-cell-specific G protein-coupled receptor
 PT polypeptides for identifying compounds that modulate taste signalling
 PT are useful in food, to modulate the sweet taste of foods or drugs -
 XX
 XX Claim 152; Page 83; 119pp; English.
 XX

CC The invention relates to mammalian taste-cell-specific G protein-coupled
 CC receptors, TIR and their corresponding cDNA molecules. Taste receptors,

CC T1R are useful for screening compounds which are used to activate or
 CC modulate chemosensory transduction, such as taste sensation. The
 CC identification and isolation of novel taste receptors and taste
 CC signalling molecules allow for new methods of chemical and genetic
 CC modulation of taste transduction pathways. The taste modulating
 CC compounds are useful in pharmaceuticals and food industries to improve
 CC the taste of a variety of consumer products, or to block undesirable
 CC tastes, e.g., in certain pharmaceuticals. T1R,s are also useful in
 CC biochemical assay for identifying tastant (T1R) ligands having binding
 CC specificity for T1R involved in taste signalling. The present sequence is
 CC human taste-cell-specific G protein-coupled receptor, hT1R1 protein.
 XX
 SQ Sequence 841 AA;

Query Match 99.98; Score 4489; DB 22; Length 841;
 Best Local Similarity 99.98; Pred. No. 0;
 Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLCCTARLVGIQLITSCWAFACHSTESSPDFTLPQDVLGALPPLHSGCLQVRHREPT 60
 DB 1 mlctarlvgllliscwafachstesspdftlpqdyllagllfpplhsgclqvrhrpevt 60
 QY 61 LCDRSSCFNEHGYHLFQAMRLGVEEINNSTALLPNTLGQYQYDVCSDSANVATLRVLS 120
 DB 61 lcdtscsfnehgylhfqamrlgveeinntallpntlgylqydvcsdsanvyatlrvls 120
 QY 121 LRGQHHTFLODDLHSTYVAIVGPDSTNRATTAALSPFLVPMISAASEFLSVKR 180
 DB 121 lrgqhhtfloodl hstytvaivgpdstnraatlaal spflvpmisyaasaelsvkr 180
 QY 181 QYPSFLRTIPMDKYQOEVEMVLLQKFGWTWISLVGSSDDYQOLGQALQENQATGGGICIA 240
 DB 181 qypsflrtipmdkyoevemvllqkfgwtwislvgssddyqolgqalqenqatgggicia 240
 QY 241 FKDIIMPFSAOYGDERMOCIMRHLAQAATVVVSSRQLARFESVVLNLTGKRWVAS 300
 DB 241 fkdimpfsaoygdermocimrhl aqaatvvvssrqlarfesvvl nltgkvwvas 300
 QY 301 EAWALSRITITVPGIORGNVLGVALQKRAVPGIKAFEEAATARADKEAPRCHKSGWSS 360
 DB 301 eawalsr ititvpgiorgnvlgvalqk ravpgikaf eeaataradkeaprchksgwss 360
 QY 361 NQLRECOAFMAHMPKIAFSMSAYNAVYAVAHGQLLGCAGASRGVYVWMO 420
 DB 361 nqlrecoafmahmpk iafsmsaynavyavah gqllgcag asrgvyvwm o 420
 QY 421 LLEQIHKVFLLHKDTVAFNDNRDPLSSYNIADWMNGPKMTFTVLGSSTPVOLNINE 480
 DB 421 lleqihkvfllhkdtvafndnr dplssyniadwmngpkmtftvlgss tpsvolnine 480
 QY 481 TKIWMGKDNQVPKSVSSDCLBEGHORVYTFHHCCFECVPCGAGTFLNKSDLYRCOPCG 540
 DB 481 tkiwmgkdnqvpk svsdclbeghorvytfh hccfecvpcg agtflnk s dlyrcopcg 540
 QY 541 KEWAPGSSQCFERTVETALREHTSWVLAANTLILLILLAGAGFAMLLDPVYVSA 600
 DB 541 kewapgssqcfertv etalrehtswv l aantlillil lagagfamll d pvyvsa 600
 QY 601 GGRVCFMLTLAGSGLSYGFEEPTRPACILRQALFALFTITFLSLCYRSPQLITIF 660
 DB 601 ggrvcfmltlags glsygf eeptrpaci l r qal f al f t i t f l s l c y r s p q l i t i f 660
 QY 661 KESTRVPTFHYAVQNHGAGLFVMISSAQLILICTLWLVVTPPLPAREYQRPFLVMEC 720
 DB 661 kestr vptfhyavq nhgaglfv m i s s a q l i l c t l w l v v t p l p a r e y q r p f l v m e c 720
 QY 721 TETNSLGFILAFYNGILISTAFACSVLGRDLPENYNKAKVTSTSLFNPSTVAFTFTA 780
 DB 721 tetnslgf ilafyng ilistafacsvlgr dlp enynk akvtstsl f npstva f t f t a 780
 QY 781 SVYDGKYLPAANMAGLSLSSGFGYFLPKCYVILCRPDLNSTEHFQASIQDYTRRCGS 840
 DB 781 svydgkylpaanm agls l s s g f g y f l p k c y v i l c r p d l n s t e h f q a s i q d y t r r c g s 840

DB 781 svydgkylpaanm agls l s s g f g y f l p k c y v i l c r p d l n s t e h f q a s i q d y t r r c g s 840
 QY 841 T 841
 DB 841 t 841

RESULT 3
 ID AAY45023
 AA45023 standard; Protein: 777 AA.
 AC AAY45023;
 DT 31-MAY-2000 (first entry)
 DE Human sensory transduction G-protein coupled receptor-B3.
 KW Human; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
 KW sensory cell; taste receptor cell; screen; taste modulator;
 KW pharmaceutical; food industry; taste topographic map; tongue.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 289
 FT N=PSDB; AAZ50745.
 FT /note= "Encoded by AC"
 XX
 PN W0200006592-A1.
 PD 10-FEB-2000.
 PE 27-JUL-1999; 99WO-US17099.
 PR 28-JUL-1998; 98US-0094465.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Zuker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;
 DR WPI: 2000-205451/18.
 DR N-PSDB; AAZ50745.
 XX
 PT New isolated sensory transduction G-protein coupled receptor, useful
 PT for developing products for use in studying and modulating the taste
 PT transduction pathway -
 PS Claim 22; Page 76; 83pp; English.
 XX
 CC The present sequence is a taste cell specific G-protein
 CC coupled receptor, GPCR-B3 which is involved in sensory transduction.
 CC This sequence was isolated from the human testis library.
 CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with
 CC lower expression in circumvallate taste receptor cells of the tongue.
 CC The present sequence is used to screen compounds that modulate sensory
 CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customise tastes. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.
 XX
 SQ Sequence 777 AA;

Query Match 90.68; Score 4068.5; DB 21; Length 777;
 Best Local Similarity 98.68; Pred. No. 0;
 Matches 767; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 64 RSCSFNEHGYHLFQAMRLGVEEINNSTALLPNTLGQYQYDVCSDSANVATLRVLSLPG 123
 DB 64 rscsfnehgylhfqamrlgveeinntallpntlgylqydvcsdsanvyatlrvls l p g 123

```

Db      1 rscsfnebhylfgamrlgveelnsta1lpnl1gylgdydcdsaanayatlrv1spg 60
Oy      124 OHHIELOGDLHYSPVLAVIGPDPSTNRATTAALISPLVPMISYAASSETLSVROPP 183
Db      61 qnhieqgdllyspcvlav1gpdscnaactaallspflv-nlisyaaasetlsvrqp 119
Oy      184 SFLRTIPNDKYOVEWVLL1QRFGWMTWISLVGSSDDYGOLGVQALENQATGOGICIAFD 243
Db      120 sflrtipndkyvevmvll1qkfgwtwislvgssddyqglvgalenglvvgiclaefdd 179
Oy      244 IMPFSQVQVDEWOCIMRLAQAATVVVVSRRQARFESVVTNTITGKVAASEM 303
Db      180 lmpfsaqvqdermqclmrllaqagavvvfssrqlarf1fesvvtlnltgkvwaseaw 239
Oy      304 ALSRHTTGVPIORIGMVGVAIOKRAVPGLKAFEEAYRADKEARPCHKSGWSSNOL 363
Db      240 alsrhtltpvgp1qrlgmvgva1qkravp1kafeeayradkearpcchksgwssn1 299
Oy      364 CRECOAFMAHTMPK1KAFSSSAYNAYRAVAVAHGLHOLLCCASGACSRGRVPMQLE 423
Db      300 crecga1mah1mpk1kafsssaynayravahg1lhc1caselcsgrvp1qle 359
Oy      424 QIHKVFHLLHKDPVAFNDNRDPLSSNITIAMDMNGPKMTFTVLGSSSTSPVOLNINERT1 483
Db      360 qihkvf1llhkdtvalndr1dp1ssyn1lawdwnp1kwt1v1gss1swp1q1ln1etk1 419
Oy      484 QHKGKDNQPKSVCSDDCLEGHORVVTGFHHCCEFCVPGAGTF1KSKSDLYRCQPCGKEE 543
Db      420 qhkgkdnqpksvcsddcleghrvvtgfhccfevcpgagtf1kkselyrcp1q1etee 479
Oy      544 WAPESQICFPRTVFLALREHTSWVLAAANTLL1LL1GTAGLEAWHLDTPVVSAGSR 603
Db      480 wapesqicfp1rvflalrehtswvl1laantll1ll1l1gtag1fawhldtpv1rsag1r 539
Oy      604 LCFIMGSLAASGS1YGFEGEPTRPAC1L1ROALFALGFTIF1SCLTVSFO1111EKS 663
Db      540 lcfimgslaa1sg1ygf1fge1trp1ac1l1rqal1tal1g1t1f1l1sc1l1v1sf1q111f1ks 599
Oy      664 TRVPFTFYHAWOVNNGAGLFVMTSSAQL1IC1L1VW1P1L1P1AREYORPH1LMECTET 723
Db      600 t1rvp1f1tyh1aw1vng1ag1lf1vmt1ssa1ql1ic1l1v1w1p1l1p1are1y1or1ph1l1m1e1c1e1t 659
Oy      724 NSLGF1LAF1LYG1L1S1T1S1F1A1C1S1Y1G1K1D1P1EN1Y1N1E1A1K1Y1T1S1L1N1F1V1S1W1A1F1T1T1A1S1Y 783
Db      660 ns1g1f1l1af1lyng1l1s1t1s1f1a1c1s1y1g1k1d1p1eny1neak1y1t1s1l1n1f1v1s1w1a1f1t1t1a1s1y 719
Oy      784 DCKVLPAAAMMAGLS1SSGFGYFLPKCYVITCRPDLNSTEFOAS1DDYPRGSGT 841
Db      720 dckvlpaaammag1l1ss1sg1fg1yfl1pkcyv1itcrp1dl1nst1efo1as1d1dypr1g1sg1t 777

```

```

RESULT 4
AAE11970
ID AAE11970 standard: Protein; 763 AA.
XX
AC AAE11970;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human novel G-protein coupled receptor (NGPCR) protein #2.
XX
KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW antisense therapy; signal transduction; behavioural disorder; obesity;
KW heartbeat rate; inflammation; immune disorder; diabetes; cancer;
KW coronary disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 29 /note= "Encoded by TTT"
FT Misc-difference 294 /note= "Encoded by RCA"
FT

```

```

XX
PN WO200172842-A2.
XX
PD 04-OCT-2001.
XX
PF 28-MAR-2001; 2001WO-US09996.
XX
PR 28-MAR-2000; 2000US-192978P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
XX
DR MPI: 2001-616474/71.
DR N-PSDB; AAD19502.
XX
PT Novel isolated polynucleotides encoding human G protein coupled
PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
PT expressed NGPCRs for diagnosis of disease, and as probes or primers
XX
PS Claim 4; Page 75-76; 80pp; English.
XX

```

The present sequence is human novel G-protein coupled receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span the cellular membrane and are involved in signal transduction after the cellular binding. The NGPCR polynucleotide sequences, are useful in diagnosis and treatment of a disease involving NGPCR, for detecting mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of disease, for screening drugs effective in treatment of symptomatic or phenotypic manifestations of perturbing the normal function of NGPCR in the body or abnormalities in the signal transduction pathway mediated by NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical trial monitoring and/or the treatment of physiological (heartbeat rate) or behavioural disorders. NGPCR is useful for identifying compounds useful in the therapeutic treatment of obesity, inflammation, immune disorders, diabetes, heart and coronary disease, metabolic disorders and cancer. Labelled NGPCR nucleotide probes can be used to screen a human genomic library which is helpful for identifying polymorphisms, determining the genomic structure of a given locus/allele and designing diagnostic tests.

Sequence 763 AA:

Query Match 90.3%; Score 4058; DB 22; Length 763;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 763; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy      79 MRLGVEEINNSYALP1PNT1LG1Q1DYVCSDSANVYAT1L1RVL1SLPGOH1E1LQ1GDL1H1SP 138
Db      1 mrlgveelnsta1lpnl1gylgdydcdsaanayatlrv1spg1h1e1lq1gdl1h1sp 60
Oy      139 TYLAV1GPDSTNRATTAALISPLVPMISYAASSETLSVROPPSFLRTIPNDKYQVET 198
Db      61 ty1av1gpd1stnra1tta1l1sp1lv1pm1isy1aas1et1sv1rk1qps1f1rt1pnd1kyq1v1et 120
Oy      199 NVYL1QKFGMTWISLVGSSDDYGOLGVQALENQATGOGICIAFKD1MPESAQVDERMQC 258
Db      121 nvyl1qkfg1wtw1slvgssddyqglvgalengat1og1c1a1f1k1d1mp1esa1q1v1der1mqc 180
Oy      259 LMRHL1QAQATVVVVSRRQARFESVVTNTITGKVAASEM1LSHRT1GVP1G1ORI 318
Db      181 lmrh1laqagatvvvvsrrqlarf1fesvvtlnltgkvwaseaw1lsrht1gvp1g1ori 240
Oy      319 GAVLGVAIOKRAVPGLKAFEEAYRADKEARPCCHKSGWSSNOLCRECOAFMAHTMPK1 378
Db      241 gav1lgva1io1kra1vp1g1kaf1ee1yar1ad1ke1ar1pc1hk1sg1w1ss1n1ol1cre1co1af1ma1ht1mp1k1 300
Oy      379 KAFSSSAYNAYRAVAVAHGLHOLLCCASGACSRGRVPMQLEOIHKVFHLLHKDPVA 438
Db      301 kafsssaynayravavahg1lhc1casgacsrgrv1pm1q1le1oi1hk1vf1h1l1hk1dp1va 360

```



```

DE Rat sensory transduction G-protein coupled receptor-B3 variant #2.
XX
KM Rat; sensory transduction: G-protein coupled receptor-B3; GPCR-B3;
KM sensory cell; taste receptor cell; screen; taste modulator;
KM pharmaceutical; food industry; taste topographic map; tongue; variant.
XX
OS Rattus sp.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 84 /note= "Wild type Glu substituted by Asp"
XX
PN M0200006592-A1.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99MO-US17099.
XX
PR 28-JUL-1998; 98US-0094465.
XX
PA (RECG ) UNIV CALIFORNIA.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Zucker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;
XX
DR WPI: 2000-205451/18.
XX
PT New isolated sensory transduction G-protein coupled receptor, useful
PT for developing products for use in studying and modulating the taste
PT transduction pathway -
XX
PS Disclosure; Page -: 83pp; English.
XX
XX The present sequence is a polymorphic variant #2 of rat taste cell
CC specific G-protein coupled receptor, GPCR-B3 which is involved in
CC sensory transduction. GPCR-B3 is specifically expressed in foliate
CC and fungiform cells, with lower expression in circumvallate taste
CC receptor cells of the tongue. The protein has an extracellular domain,
CC seven transmembrane domains and an intracellular domain.
CC The GPCR-B3 sequence is used to screen compounds that modulate sensory
CC signalling in taste cells, especially taste modulators useful in
CC pharmaceutical and food industries to customise taste. The sequence
CC can also be used as probe for identifying taste cells and
CC subsets of taste receptor cells such as foliate, fungiform and
CC circumvallate. Such probes are also useful to generate taste
CC topographic maps that elucidate the relationship between the taste
CC cells of the tongue and sensory neurons leading to taste centres
CC in the brain.
CC Note: The present sequence is not given in the specification but is
CC derived from rat GPCR-B3 sequence shown in page 75 (AAV45021).
XX
XX Sequence 840 AA;
SQ

```

```

Query Match 74.7%; Score 3358.5; DB 21; Length 840;
Best Local Similarity 74.0%; Pred. No. 0;
Matches 622; Conservative 82; Mismatches 136; Indels 1; Gaps 1;

```

```

Db 180 kfpstlrvpsdrbqvewmvgllqsfgwvwlslqsydydgqlygqaleelavprglcva 239
QY 241 FKDI MPESAQYGDERMOCIMRHLAQAATVYVVVSSROLARFESVLTNITGKWVAS 300
Db 240 fkdlyvfaarvgydprmgsmqhlqartlvvsnrlavfrsvlanltgkwvas 299
QY 301 EAMALSRHITGVPGIORTIGMVLGVAIORKRAVPGKAFPEAVARADKEAPRCHKSMCSS 360
Db 300 edwalstylvtsqlygqlygvtlvgyavqgrvpyglkeifeesyvrvatapsacpewscst 359
QY 361 NQLRCCQAFMAHMPKLFKAFSSMSAYNAVAVAHGLHQLGACASGSRGRVYPMQ 420
Db 360 nqlcrechlttrmptlrgatfmsaayrvyavayvalngllqgtseleisrgrvypwq 419
QY 421 LLEQIHKAVHLLHKDTYAFNDNRPLSSYNIANDMNGPKMTFYVLGSGSTSPQVLNNE 480
Db 420 llqglykvnflilhentvafidngdtlgydlawdngpewtfeilgsaslsphldlnk 479
QY 481 TKIOWHGKDNQVPSKSCSDCLEGHQRVVTGPHHCCFECVPCGAGTFLNKSDLYRCOPCG 540
Db 480 tkigwghkndgvpsvctcdclaghnrvvsgshnccfcpceagclfmseilhcyqcg 539
QY 541 KEENAPESGQTCFPTVYVFLALREHTSWLLAANTLTLGLTGAAGLRPMHLDPVVRSA 600
Db 540 teewapkestctcfprtvelflawhepislvlantllllllvgtagltawhfhprvrsa 599
QY 601 GGRICFTMLGSLAAGSGSLYGFEGEPTRPACILRQALFALGFTTFLSCLTVRSROLIIF 660
Db 600 ggricftmlgslavagscsfyfsgfeptrpaciilrpilslgfafilscltrsfqlvlf 659
QY 661 KESTKVPTEFYHAMVONHAGLFVMISAAOLICTWLVVMTPPAPAREGRPHVMMEC 720
Db 660 kfstkvpftrtwagngaglfvisscvhllclctwlvmtpprptcyqifphvlliec 719
QY 721 TERNSLGFIALFLYNGLLSISAFACSYLGKDLPENYNAKCVTSLSLFNYSWIAFFTTA 780
Db 720 tevnsgfliaflthllslstfscysylgkelpenyneakcvtslllnfswiaffma 779
QY 781 SYDGKTYIPANMMAGLSLSSGFGYFLPKCYVILCPDUNSTHFPASITODTYRRGCS 840
Db 780 slqgsyilpavnavlagltlsgisgyflpkcyvillcrpelnthetqasldqylrrcgt 839
QY 841 T 841
Db 840 t 840

```

```

RESULT 8
ID AAV45028 standard; Protein: 840 AA.
AAV45028;
31-MAY-2000 (first entry)

```

```

DE Rat sensory transduction G-protein coupled receptor-B3 variant #3.
XX
KM Rat; sensory transduction: G-protein coupled receptor-B3; GPCR-B3;
KM sensory cell; taste receptor cell; screen; taste modulator;
KM pharmaceutical; food industry; taste topographic map; tongue; variant.
XX
OS Rattus sp.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 90 /note= "Wild type Ala substituted by Gly"
XX
PN M0200006592-A1.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99MO-US17099.

```


XX 28-JUL-1998; 9805-0094465.
 PR (REGC) UNIV CALIFORNIA.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Zuker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;
 DR WPI: 2000-205451/18.
 XX
 PT New isolated sensory transduction G-protein coupled receptor. useful
 PI for developing products for use in studying and modulating the taste
 PT transduction pathway -
 XX
 PS Disclosure; Page -: 83pp; English.

CC The present sequence is a polymorphic variant #3 of rat taste cell
 CC specific G-protein coupled receptor, GPCR-B3 which is involved in
 CC sensory transduction. GPCR-B3 is specifically expressed in foliate
 CC and fungiform cells, with lower expression in circumvallate taste
 CC receptor cells of the tongue. The protein has an extracellular domain,
 CC seven transmembrane domains and an intracellular domain.
 CC The GPCR-B3 sequence is used to screen compounds that modulate sensory
 CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customise taste. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.
 CC Note: The present sequence is not given in the specification but is
 CC derived from rat GPCR-B3 sequence shown in page 75 (AAV45021).
 CC
 XX
 SQ Sequence 840 AA;

Query Match 74.7%; Score 3357.5; DB 21; Length 840;
 Best Local Similarity 74.0%; Pred. No. 0;
 Matches 622; Conservative 81; Mismatches 137; Indels 1; Gaps 1;

QY 1 MLCTRRLVGLQLLISCCAFACHSTRESSPDITLPDYLALGFLPHSGCLQYRHRPEYT 60
 DB 1 mlfwaahl-lslqlvcafwascqrlesspgfslpgdflglfshgclqlvrhplvt 59
 QY 61 LCDRSCSFENEHGYHLFOARLGEVIEINNSTALLPNTTLGQLYDYCSDANVATLRYLS 120
 DB 60 scdrpsdfngghyhlfganrfveelnnsgyllpntllgelydvcsesanyatrlvla 119
 QY 121 LFGQHIEIQLGDLHSPYLAIVIGPDSTNRATTAALISPELVPMISYAASETLSVKR 180
 DB 120 lqgprhlelqkdlrhnskvaflgpdntdhavtaallgpfimpvysaasvlsakr 179
 QY 181 QYPSFLRTIPNDKYQYEVNVLLOKFGWWSISNGSSDYQGLQVALENQATGOCICIA 240
 DB 180 ktpstlrvpsdhrhvevnmqllqsfqwwslsgysgqglqyqvalaelavpricva 239
 QY 241 FFDIMPFSAOVDERMOCCLRHIAQAGATVAVVVFSSROLARFEFEEVTVNTLTKWVMS 300
 DB 240 fkdipfsarvgdprmmgmhlaqartlvvfvfshrlharffrsvvanltgkwvms 299
 QY 301 EWMALSRHITGVPGIORIGMVLGVALQKRAVPGLKAFEEAVYARADKADRPCHKSGWSS 360
 DB 300 edwaistylstvgigltvgavqgrvpgljkefesyavraaapacpegswcst 359
 QY 361 NOLCRECAFMAMHTPKLKAFGSSAXNMYRAVYANAHLHLQLGASACSGRYPYPMQ 420
 DB 360 nqlcrechftrcmptlgaifmsaayryveayavahhqlhqltseicstgpyvpmq 419
 QY 421 LLEQIHKVFLAKDPAVANDNDPLSSYNIAMDNNGKMTFTVYGGSTWSPVOLINIB 480
 DB 420 llqgylkvfllhentvaldngdtlgyddllawdngpewfcltsgsslsipvhlidnk 479

QY 481 TKIOWHGKDNQPKSVYSSDCLEGHORVYTGHHCCPEVCPCAGTFLKSDLYRKOPCG 540
 DB 480 tkqyhgknqpyvsvctdclaghnrvvvgshhccfecvpcceagtlflmsehlhlcpcg 539
 QY 541 KEEMAPESGSCPPRTVAVFLALREHNSWVLANTLILLIGTAGLPAWHIDTPVRSR 600
 DB 540 teewapkestcprtrveflawhepislvlantlllllyagtaglfawhfpvrvsa 599
 QY 601 GGRICFLMLGSLAAGSGSLYGFEGEPTRPACILQALFALGFTIFLSCLTVRSFQILIIF 660
 DB 600 ggricflmlgslavagscsfysffgeptpacilrqlgflgafiflsciltsfqlvlf 659
 QY 661 KSTKVPTFYHAMVONHGALEFYMISSAOLICLTWLVVTPLPAREYQRPPLVMEC 720
 DB 660 kstkvptfyrwagngaglfvlyvsvthlllcltwlvvtprepreyqrfpvlvlec 719
 QY 721 TEHNSGFLIAPLYNGLLISAFACGYLKDIPENNEKCVTFSLFNPVSLIAETFA 780
 DB 720 tevnsgfliafthlllslstvcslgkelpenyneakcvfslhntvsliaetfa 779
 QY 781 SVYDGKYLPAANMAGLSLSGFGYFLPKCVIILCRPDLNSTEHFOASIDYPRGCS 840
 DB 780 slvgsgylpavnvlaglittlsgfsgyflpkcviilcrpelnstehfsgsidytrrgt 839
 QY 841 T 841
 DB 840 t 840

RESULT 9

AAV45022
 ID AAV45022 standard; Protein; 842 AA.

AC AAV45022;

DT 31-MAY-2000 (first entry)

DE Mouse sensory transduction G-protein coupled receptor-B3.

XX Mouse; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;

KW sensory cell; taste receptor cell; screen; taste modulator;

XX pharmaceutical; food industry; taste topographic map; tongue.

OS Mus sp.

PN WO200006592-A1.

PD 10-FEB-2000.

PF 27-JUL-1999; 99WO-US17099.

PR 28-JUL-1998; 98US-0094465.

PA (REGC) UNIV CALIFORNIA.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Zuker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;

DR WPI: 2000-205451/18.

DR N-PSDB; AAV50744.

PS Claim 22; Page 75; 83pp; English.

CC The present sequence is a taste cell specific G-protein

CC coupled receptor, GPCR-B3 which is involved in sensory transduction.

CC This sequence was obtained from mouse circumvallate and foliate papillae.

CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with

CC lower expression in circumvallate taste receptor cells of the tongue.

CC The present sequence is used to screen compounds that modulate sensory

CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customise taste. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.

XX Sequence 842 AA:

Query Match 74.5%; Score 3348.5; DB 21; Length 842;

Best Local Similarity 73.5%; Pred. No. 0;

Matches 619; Conservative 87; Mismatches 135; Indels 1; Gaps 1;

QY 1 MLLCTAR LVGLQLISCCMAFACHSTSSDPFTLPGLYLLAGFLPHSGCLQVHRPEV 59
 DB 1 mlfaaahlllslqavaycwfscqfrespsqfslpgdflilagflshadclqvrhplv 60
 QY 60 TLCDRSCSFENHGVLFOAMRLGVEINNSTALLPNTLGLQVLYVCSDSANVATLRLV 119
 DB 61 tscdtsdfngnyhllqamfveelnstallpnlclgelyvcsessnyactltvp 120
 QY 120 SLPGQHILTELOGLDHLNPTVLAVIGPSTNRATTAALSPFLPMISYAASSETLSVK 179
 DB 121 aggttghlemgrdlrnhsckvvaaligpndchavtaaalspflmlyssvllsgk 180
 QY 180 ROYSEFLRTINDKYQVETWVLLQKFGWTWISLVGSSDDYGLQGLVLENDATQOGICI 239
 DB 181 rkfpstftrlpdsdyqvavivrlqsfqvwslvsgdyqqlvgqaleatprglcv 240
 QY 240 AFKIDMPASQVGDDEMRCLMRHLAQAATVYVVFSSROLARVPESVTLTKKVVWA 299
 DB 241 afkdvpslqagadpdmrgmmlrlartctvvvfnshlgvtrsvlanllgkvwa 300
 QY 300 SEAMALSHRTGVPGIORIGVAVIQRAVPGIKAFEEAYARADKEAPRCHKSMCS 359
 DB 301 sedwaistyltnvpglqglvlgvavlgqvpglkefeesyvqvmgaptcpgswcg 360
 QY 360 SNOUCRECAQAMHTMPKLFKAFSSSAYNATRAYAVAHGLHQLGCSGACSRKRVYFW 419
 DB 361 tnglcrechaitwmpelgaifmsaayvayvabhqllgicstgarqpyvpw 420
 QY 420 QLEQIHVHFLHNDYAFNDNRDPLSSYNIADWNGPKMTFVLSSFTWSPQOLNIN 479
 DB 421 qlldqiyvnlflhkktyafadkdgdpdygyliawdmngpewtlevisasalspvhldn 480
 QY 480 ETKIQMHGKDNQVPSVSSDCLLEGHORVVTGFHHCCECVPCGATFLNKSNDLYRCPC 539
 DB 481 ktklqwhgknqyvsctrcldqeghrlvmgshhccfcmpeceagftlnselhtcpgc 540
 QY 540 GKREKAPRGSGOTCFPRYVFLALRHTSWLLAANTLLLLLLGTAGLFAHMLDTPVRS 599
 DB 541 gteewapessactsrteflgwhpelslvllaentlllllgtaglfawrlhpyvrs 600
 QY 600 AGGRCLFMLSLAAGSGLGFCFCEPRPACILROALFALGFTFTLSCLTVRSQOLITI 659
 DB 601 aggrclfmisglvagscslysfqkpvpcalllqplfsigfaiflscltrstqlvll 660
 QY 660 FKSTKVPFTHAVQNHAGLFWMISSAOLICTLWLVVTPPLPAREYORPHLVMLE 719
 DB 661 fkfstkvpfthwqngagifvsvstvhllfcltvlamtpptretyqfplvllle 720
 QY 720 CTETNSLGFILAFYNGNLISAFACSTLGDLPENYNEAKCVTSLLFNFSYIAFTTT 779
 DB 721 ctetnsyglfvaafhnlisistfvcslgkelpenyneakcvtslllhfvswiaftfm 780
 QY 780 ASVYDGYFLPAPNMMAGLSLSSGCGYFLPKCYITLCRPDINSNEHROASIODYTRRG 839
 DB 781 sslyggsylpavnvlaagiatlsgtsgyflpkyvyltcrpelntenhqasldqyltrcg 840
 QY 840 ST 841

DB 841 tt 842

RESULT 10
 ID AAE11971 standard; Protein; 366 AA.

XX AAE11971;

DE 18-DEC-2001 (first entry)

XX Human novel G-protein coupled receptor (NGPCR) protein #3.

KW Human: G-protein coupled receptor; GPCR; gene therapy; drug screening;
 KW antisense-therapy; signal transduction; behavioural disorder; obesity;
 KW heartbeat rate; inflammation; immune disorder; diabetes; cancer;
 KW coronary disease.

OS Homo sapiens.

EH Key Location/Qualifiers

FT Misc-difference 174 /note= "Encoded by RCA"

XX MO200172842-A2.

XX 04-OCT-2001.

XX 28-MAR-2001; 2001WO-US09996.

XX 28-MAR-2000; 2000US-192978P.

XX (LEXI-) LEXICON GENETICS INC.

XX Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;
 PI Zambrowicz B, Sands AT;

XX WPI: 2001-616474/71.

XX N-PSDB: AAD19503.

XX Novel isolated polynucleotides encoding human G protein coupled
 PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 PT expressed NGPCRs for diagnosis of disease, and as probes or primers -

XX Claim 5; Page 77-78; 80pp; English.

XX The present sequence is human novel G-protein coupled receptor
 CC (NGPCR) protein. NGPCRs are transmembrane proteins that span
 CC the cellular membrane and are involved in signal transduction after
 CC ligand binding. The NGPCR polynucleotide sequences, are useful
 CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heartbeat rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.

XX Sequence 366 AA;

Query Match 40.1%; Score 1803; DB 22; Length 366;
 Best Local Similarity 99.4%; Pred. No. 4e-174;
 Matches 334; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

[illegible]

Search completed: June 11, 2002, 10:17:41
Job time: 3813 sec

```

|||||
DB 262 CTCCTCCGCGGACCTCAAGACGACAGGACCTCCAGGCTCCATTCAGGACCTACACG 203
|||||
QY 2506 aggcgtcgcgcctcacctga 2526
|||||
DB 202 AGGCGCTGCGGCTCCACCTGA 182
|||||

RESULT 3
AA853967/c 496 bp mRNA linear EST 31-DEC-1998
LOCUS 31j01.01 Soares_testis.NHT Homo sapiens cDNA clone IMAGE:1393866
DEFINITION 3' splice to 5N:CA5R.BAT P48442 EXTRACELLULAR CALCIUM-SENSING
RECEPTOR PRECURSOR ;, mRNA sequence.
ACCESSION AA853967
VERSION AA853967.1 GI:2941505
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 496)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rt@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the J.M.A.G.E. Consortium/LNLW at:
www-bio.llnl.gov/dbfp/image/image.html

Possible reversed clone: similarity on wrong strand
Insert length: 878 Std Error: 0.00
Seq primer: -40ml3 fwd. ER from Amersham
High quality sequence stop: 223.
Location/Qualifiers
1. 496
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1393866"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAAGTGGGCGCGCCCAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 102 a 143 c 148 g 100 t 1 others
ORIGIN

Query Match 11.0%; Score 278.6; DB 9; Length 496;
Best Local Similarity 95.6%; Pred. No. 2.8e-55;
Matches 306; Conservative 1; Mismatches 11; Indels 2; Gaps 2;
QY 2207 gccctctccatcagctccttcgctcgaagctacgtgtaaggaagctcgcagaagct 2266
|||||
DB 495 GCGCTCTCCATCATGATGCTTCCCTGACAGTACTGGGTAAAGACTTGGCCAGAGACT 436
|||||
QY 2267 acaacagagcaaatggtacacctcagcgtccttcaactcgtctcgtgacgcct 2326
|||||
DB 435 ACAACAGAGGCCAAATGTCTCACTTCAAGCTGCTCTCAATTCGTCTGATCGCCT 376
|||||

```

```

QY 2327 tcttaccacagcgcagcgtctacagcagcgaagctactcgtcgcgcacaaatgatgctg 2386
|||||
DB 375 TCTTCACCAAGGCGGACGTCCTTACAGACGCAAGTA-ANCCCTTGCGGCCAACATGATGGCT- 316
|||||
QY 2387 ggcctagacagcctgagcagcgcgtcgtggtatcttcctgctaaagctacgtatcc 2446
|||||
DB 317 GCGTACGACGCTTGAGCAGCGGCTTCGGTGGGTATTTCCTGCTTAAGTCTACGTGATCC 258
|||||
QY 2447 tctgcgcgcacagcctcaacagcagcagcactccagcgtcctcaatgaagctacaga 2506
|||||
DB 257 TCTGCCGCCACAGACCTCAACAGCAGACGACCTTCAGGCGCTTCATTCAGGACTACAGA 198
|||||
QY 2507 ggcgtcgcgcctcacctga 2526
|||||
DB 197 GCGCTGCGGCTCCACCTGA 178
|||||

RESULT 4
BB618551/c 666 bp mRNA linear EST 26-OCT-2001
LOCUS BB618551 RIKEN full-length enriched, 8 days embryo Mus musculus
DEFINITION cDNA clone 5730408M21 5', mRNA sequence.
ACCESSION BB618551
VERSION BB618551.1 GI:16458094
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 666)
AUTHORS Arakawa,T., Carlini,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Komoto,H., Kouda
,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Susehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carlini,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Komno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
vagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Iwata,M., Ohara,E.,
Watanabe,M., Yonekura,Y., Ishikawa,T., Ozawa,Y., Tanaka,T., Matsura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carlini,P., Sugahara
,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shingawa,A., Saito,T., Kiyosawa,H., Yamakata,I., Atzawa
,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

```

FEATURES	e mouse	tissues.	Location/Qualifiers
SOURCE	1..666		/organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="5730408M21" /clone.lib="RIKEN full-length enriched, 8 days embryo" /sex="mixed" /dev_stage="8 days embryo" /lab_host="DH10B" /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15', GAGGAGAGAGATTCGATTAATTAATTAATTAATTCACCCCCCCCCC 3']. cDNA was prepared by using trehalose thermoactivated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 100.0 Second strand cDNA was prepared with the primer adapter of sequence 15', GAGGAGAGATTCGATTAATTAATTAATTAATTCACCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pluscript KS(+) after bulk exsition from LambdaFLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
BASE COUNT	152 a	189 c	181 g 144 t
ORIGIN			
Query Match	9.6%;	Score 242.2;	DB 9; Length 666;
Best Local Similarity	79.0%;	Pred. No. 1.3e-46;	
Matches 312;	Conservative	1;	Mismatches 79; Indels 3; Gaps 2.
Oy	2132	gcttcccccatctggtatgctgctgagtcacagagaccacccctggtcctcaactg	2191
Db	643	gctttcccccctatcggttaattcgtt-AGTGCACAAAGGCTCATAT-CTTGGGGTTCTTGGGG	587
Oy	2192	cctctctcacaaatggcctcctctcatcatagtcgcttcttcgagacacccctgggttaag	2251
Db	586	gctttggcacacacaaatcctctctctcatcatcagacaccttctgacagctacctggggaag	527
Oy	2252	acttgcacagaaactacacagagacgaatgcttcaactcctcagcctgctcctcaactcg	2311
Db	526	AACTCGGGGAGAACTATAACGAAGCAAAATGTCACCTTCAGCGCTCCACTTGG	467
Oy	2312	tgtcctgtagctgcctctctcacacgagccagcgtctcagacgagcaagtaactgctgcg	2371
Db	466	TATCCTGTGATCGCTTTCCTACCACTGTCAGGATTTTACCGAGGCGACTACCGCG	407
Oy	2372	ccaacatgtagctgctgagcagcagcctgagagagagcctggtgtgtattcttcgccta	2431
Db	406	TCAATGTGTGGCAAGGGCGGCGCCACTCTGAGTGGGGGCTTCAGCGGCTATTTCTCCCTA	347
Oy	2432	agtgctacgtgtagctctctgacgcccagacactcaacagacagagacactcagagcctca	2491
Db	346	AATGTAGTGATTTCTTGCGCGTCCAGAACTCAACACACACAAACACTTTCAGGCTCCA	287
Oy	2492	ttcaggaactacagagcgctgcggtcctccactga	2526
Db	286	TCCAGGACTACACGAGGCGCTCGGCGACTACTGA	252
RESULT	5		
LOCUS	AI390993	525 bp	mRNA linear EST 15-MAR-2000
DEFINITION	mb98d01.y1	Scores mouse p33MPF19.5	Mus musculus cDNA clone IMAGE:337441 5' similar to SM-CAR_HUMAN P4180 EXTRACELLULAR CACIUM-SENSING RECEPTOR PRECURSOR ; mRNA sequence.
ACCESSION	AI390993		
VERSION	AI390993.1	GI:4217000	

KEYWORDS EST.
SOURCE Mus mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 525)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theisling,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R., and Wilson,R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the correct orientation)
MGI:218841
Seq primer: -40RP from Gibco
High quality sequence stop: 481
POLYA-No.
Location/Qualifiers
1. .525

FEATURES	SOURCE	Location/Qualifiers
		1..525
		/organism="Mus musculus"
		/db_xref="taxon:10090"
		/clone="IMAGE:337441"
		/clone_lib="Soares mouse p3NMF19.5"
		/dev_stage="19.5 dpc total fetus"
		/lab_host="DH10B (ampicillin resistant)"
		/note="Vector: pRT73D (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGGCGCATTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pRT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldi. RNA was kindly provided by Dr. Mitsuhiro Ko (Wayne State University)."
BASE COUNT	104 a	179 c 102 g 140 t
ORIGIN		
Query Match	8.3%	Score 210; DB 9; Length 525;
Best Local Similarity	63.9%	Pred. No. 4.9e-39;
Matches 399; Conservative 1;	Mismatches 111;	Indels 113; Gaps 1;
QY 1854	cctctatgctctcttctgggaacccaagaagcgtcgtctgtcgaagcaagcctctt	1913
DB 15	CCACCAAGGAGTACCAAGCGCTTCCCCCATCTGGTGATTTCTTGCTGCTACGCCCTCTT	74
QY 1914	tgcccttggttcaacacatctctctgtctctgtcgtgaagttcgtaattccaaactaatcat	1973
DB 75	TTTCTCTGGGTTTGGCATTTTCCCTCTCTCTGTCGACAAATCCGCTCTTCCAACTGGTCAT	134
QY 1974	catcttcaagttcttccaccaaggttaactcaaatcttaccacagcgtggtggtccaaacacagg	2033
DB 135	CATCTTCAAGTTTTCTTCAACCAAGTACCCACATTTTACACACTTGGGCCCAAAACCATAGG	194
QY 2034	tgctggacctgtttgtaagatcaagtcagcgagccacagctgtactatcttcaacttgct	2093
DB 195	TGCCGGAAATATTTCGTCATTTGTCAGTCCACGCTCATTTGTTTCCCTGTGTCTCAC-----	248
QY 2094	ggtggtgtggaaccactgctctgtagggaataaccagcgcttcccacatctgtgtatgct	2153

Db 249 ----- 248

Qy 2154 tgagtgcacagagacaaactccctgggtctatactgctctctctacatgagctct 2213
 Db 249 ----- ACAACATCTCTCT 261

Qy 2214 ctccatcagtccttctgctcagctacgtgtaagagcttgcacagaaactacaaga 2273
 Db 262 CTCATCAGCAGCCTTTGCTGCTGAGTACCTGGTAAAGAACTGCGGAACTATAAGCA 321

Qy 2274 ggcacaatgyltcactcagctcgtctcttaactcgtctcgtgagctgctcttac 2333
 Db 322 AGCCAAATGTGTACCTCAGCTCAGCTGCTCCTCCTACTTCCTATCCTGAGTCTTCTTAC 381

Qy 2334 cagggcagaggtctacagacagaaactacgtcgtcgcacaaatgagctgagctgag 2393
 Db 382 CATGTCACGACATTAACGAGGAGCTACTACCCGCGGTCAATGCTGAGGAGGCTGCG 441

Qy 2394 cagcctgagcagcgtctcgtgtggtatctctcgtcagtgagctacgtcgtcgcg 2453
 Db 442 CACTCTGAGTGGCGGCTTCAGCGGCTATTCTCCTCCCTAAATGCTACGTATCTCTGCG 501

Qy 2454 ccagacctcaacagcagacaga 2477
 Db 502 TCAGAACTCAACACAGACAGA 525

RESULT 6
 LOCUS CNS033X/c 1014 bp DNA linear GSS 15-MAY-2000
 DEFINITION Tetradon nigriviridis genome survey sequence T7 end of clone 209N08 of library G from Tetradon nigriviridis, genomic survey sequence.
 accession AL226735
 VERSION AL226735.1 GI:7885667
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetradon nigriviridis.
 ORGANISM Tetradon nigriviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetradon.
 1 (bases 1 to 1014)
 Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigriviridis
 Unpublished
 2 (bases 1 to 1014)
 Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using Tetradon nigriviridis DNA sequence
 Unpublished
 3 (bases 1 to 1014)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigriviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetradon>.
 FEATURES
 source 1..1014
 /organism="Tetradon nigriviridis"
 /db_xref="taxon:99883"
 /clone="209N08"
 /clone_1bp="G"
 /note="Genoscope sequence ID : COAG209D04LP1-end : T7"
 BASE COUNT 283 a 209 c 322 g 196 t 4 others

Query Match 8.2% Score 208; DB 12; Length 1014;
 Best Local Similarity 53.2%; Pred. No. 2e-38;
 Matches 478; Conservative 3; Mismatches 414; Indels 4; Gaps 2;

Qy 1594 gaccctcaagatgacagcctgtgtggaagaagagtggtgacactgaggaagaccagc 1653
 Db 896 GATCCCTACAGCTGCTCCTCCCTGTAAGAAGACACAGCTGTCCACAGCTGGAAAGACCTCC 837

Qy 1654 tgcctccgagcagctggtgtgttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1712
 Db 836 TKTCTCCACGAGAGGTGAGAGGTGCGGTTCACAGACCCCGCGGTGAGTATCTCT 777

Qy 1713 ggcagctaacacgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1772
 Db 776 GCGCCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717

Qy 1773 cctagacacccctgt 1832
 Db 716 CTACAAACAGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657

Qy 1833 cctggcagcaggtgagtgagcagcctctatgctctctgtgtgtgtgtgtgtgtgtgt 1892
 Db 656 CCTCAGCCTCTGTGATCAGTACAGCGCTCTCTTACTTGAAGACCGATGAAGCCTTTTG 597

Qy 1893 ctgtctagcagcagcctctgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1952
 Db 596 TATCTTGAAGTTTATGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 537

Qy 1953 tgcgtcatcacaatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 2012
 Db 536 CCGCTCTTTCAGATGATCATATTTTAAATATACCGGCAAGTTCCTCCGAGTCCACAG 477

Qy 2013 cgcctgtgtccaaacacacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2072
 Db 476 CTGGTGATGAAATGACACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417

Qy 2073 gctatcgtctaaactgt 2132
 Db 416 AGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 357

Qy 2133 ctcccccatctgt 2192
 Db 356 CTACCCAGCAAAATCAATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

Qy 2193 ctctctcaaatgt 2252
 Db 299 ATTTTGTGCTTCTTATGTTGTGATCCTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 240

Qy 2253 ctggcagagaaactacaacagagcaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2312
 Db 239 CTTCCCGAAGAACTACAAAGAGGCGCAAGACCTAATCTTCTTCTTCTTCTTCTTCTTCTTCT 180

Qy 2313 gtctgtatcgcctctctctcaacagcagcagcagcagcagcagcagcagcagcagcagc 2372
 Db 179 CACCTGATCATATTTGGCACCGGCTTCATGCTACACAGGCAATATCATCACACCTT 120

Qy 2373 caacatgagctgt 2432
 Db 119 CAACCCCTGCGCGGTGCTCTCCAGCGCTTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 60

Qy 2433 gtgtcagtgatcctctgt 2491
 Db 59 ATGTTACATTAATCATCTTTCAGCTTCACAGAACACCAAGAAAGTACTTCCAACTTCCA 1

RESULT 7
 LOCUS A1415100/c 421 bp mRNA linear EST 09-FEB-1999
 DEFINITION mb98d01.x1 Soares mouse p13NF19.5 mus musculus cDNA clone IMAGE:337441 3' similar to SW:CAIR_EAT P48442 EXTRACELLULAR CALCIUM-SENSING RECEPTOR PRECURSOR ;, mRNA sequence.

Db 181 ACGAGCGCTGCGGCACTACTGA 158

```
On:    365 ++++++awt+cact+ccactcc
Matches 219; Conservative 1; Mismatch
```

QY 2265 ctacaacgaagccaatatggtacacttcagcctgtctctcaacttcgtgtcctgattcgc 232

); Site-1: XhoI; Site-2: EcoRI; cDNA oligo df-primed
[5'-(GA)10-ACTAGTCGAGTGTGTTT-3'] and directionally
cloned using 5' linkers 5'-AATCGCAGCAG-3' and
5'-CTCGGCGC-3'. Size selection of >400bp material gives
average insert size ranging from 1-2 kb. Library was mass
excised (from Lambda-UniZAP-XR) and resulting
single-stranded phagemids were prepped and transformed
into DH10B. Library contains 98 recombinants.
References: J. Androl. 20:635-639 and Gene 25:263-269.
Library constructed and donated by J. McCarrey, Ph.D.
(Southwest Foundation for Biomedical Research, Dept. of
Genetics); excision done by E.M. Eddy, Ph.D. (National
Institutes of Health, National Institute of Environmental
Health Sciences). Original lambda-based library is
available through ATCC, catalog #63422."

BASE COUNT 63 a 67 c 69 g 63 t
ORIGIN

Query Match 7.0%; Score 177.4; DB 10; Length 262;
Best Local Similarity 80.3%; Pred. No. 1.8e-31;
Matches 208; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 1379 agtggacccttaacggtcctcgtgtccacatggtcctcaagtcagcttaataatg 1438
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3 AATGACCTTTGAGTGATGATGTTGCTCCTCACTGCTCAGTCAATCAAGACATAAATA 62
QY 1439 agaccaaattccagtgacagcagaaagacacaggtgtccctgaagctgtgtgtccagc 1498
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 63 AGACAAATATCCAGTGGCAGCAGGAGAAACATAGAGTGTGTCTGATGTATCCAGG 122
QY 1499 actgtctgaagggcaccagcagagtggttaacaggttccatcagctgtctgttaagtg 1558
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 123 ACTGTCTCAAGAGGACACACAGAGTGTGATGGTCCACACCACTGCTGCTGAGATGCA 182
QY 1559 tggcctgttgggttggtgaccttctcaacaaggtgaacctctcaatgccaactgtgtg 1618
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 183 TGCCCTGTGAAGCTGGACATTTGTCAACACAGTAGAGCTTCACTGACCTGGCAGCTTGTG 242
QY 1619 ggaagaagaagtgagcacc 1637
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 243 GAACATATAGATTGGGCC 261

RESULT 11
A2451238 502 bp DNA linear GSS 04-OCT-2000
LOCUS 1M0250F04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0250F04 R, DNA sequence.
ACCESSION A2451238
VERSION A2451238.1 GI:10606843
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 502)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

Plate: 0250 row: F column: 04
Seq primer: CACACAGAAACGCTATGACC
Class: plasmid ends
High quality sequence stop: 502.
Location/Qualifiers
1..502

FEATURES
source

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0250F04"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42uv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g14732141g14732141g14732141), a copy number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 103 a 139 c 151 g 108 t 1 others
ORIGIN

Query Match 6.6%; Score 166.4; DB 12; Length 502;
Best Local Similarity 82.3%; Pred. No. 9.9e-29;
Matches 191; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2295 cctgcctcaactgcttctcgtatgctccttcaccacagcagcgtctagcagcg 2354
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 CCGCTCTCCACCTTGATTCCTGGATCGCTTCTTCCACATGTCCAGATTACCAAGG 60
QY 2355 caagtaacctgctcgccacaacatgatagtgctgagcagcagcgtctgagc 2414
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 CAGCTACGTACACCGCGGTCAATGTCTGGCAGAGGCTGGCCACTGTAGTGGCGCTTCAG 120
QY 2415 tgggtatttctgctcaagtgtcagtgatctcttcgcgcagcagcactaaacagcaca 2474
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 CGGCTATTACCTCCCTTAATGCTACGTGATTTCTCCCTCCAGAACTCAACAAACAGCA 180
QY 2475 gcaactcagagcctccatcagacacagcagcagcgtcgcagcactga 2526
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 ACACTTTACGGCTCCATCCATCAGACATACAGAGCGCTCGGCACACTCTGA 232

RESULT 12
BM439858 612 bp mRNA linear EST 01-FEB-2002
LOCUS pgrin.ph001.h19 Normalized Chicken Reproductive Tract cDNA Library
DEFINITION (pgrin) gallus gallus cDNA clone pgrin.ph001.h19 5' similar to
gblABD18069.1 (AF127389) putative taste receptor TR1 (Rattus
norvegicus), mRNA sequence.
ACCESSION BM439858
VERSION BM439858.1 GI:18470633
KEYWORDS EST.
SOURCE Chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 612)

Db 321 CCAAAACATGGGCGCCGACACATTTTCATCCATCTGCGCCACATCTGTGCTC 262
 Oy 2081 gctaacctggtgctggtgagaccccaactgctgctgagataccagcctccccc 2140
 Db 261 CTGT-CTGGGCGTGAACCGCTGGGCCCCCGACGCTCCARRATCTTATTTTCACGA 203
 Oy 2141 atctgtgagctggtgagacagagacacacacccctggtgctcatctgctcctc 2200
 Db 202 ACAGCATCGMGCTGRAGMGACAGACACCTCTGCGCCGTTCTCTTGAGAGACTGTGA 143
 Oy 2201 acaatgctcctctcctcagctgctgctgctgctgctgctgctgctgctgctg 2260
 Db 142 AGTCATGTCTGAGCGCGCTGCTGCTTTTCAGCMAATGAGGCAAAARACTCTCCG 83
 Oy 2261 agaatcaacagagagcaaatggtgctgctgctgctgctgctgctgctgctgctg 2220
 Db 82 CCAAGTACAGAGGCGCAAGTGTGCTACCTTCAGCCTGATGTGAATATATCTCTGGA 23
 Oy 2321 tgcctctctcacc 2334
 Db 22 TTAGCTTCTTAC 9

RESULT 14
 W18663 247 bp mRNA linear EST 10-SEP-1996
 LOCUS mb98d01.r1 Soares mouse p3NMf19.5 Mus musculus CDNA clone
 DEFINITION IMAGE:337441 5', mRNA sequence.
 ACCESSION W18663
 VERSION W18663.1 GI:1294371
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 247)
 Maeda, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Gelsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through INMIL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:218841

Seq primer: mob.REGA+ET
 High quality sequence stop: 225.
 Location/Qualifiers

FEATURES
 source
 1..247

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:337441"
 /clone_lib="Soares mouse p3NMf19.5"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Vector: pRT73d (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTACCAATCTGAGTGGAGCGCGCCGATTTTCTTTTCTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pRT73 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento

Soares and M.Fatima Bernaldo. RNA was kindly provided by
 Dr. Minoru Ko (Wayne State University)."
 BASE COUNT 44 a 89 c 42 g 72 t
 ORIGIN

Query Match 5.2%; Score 131.6; DB 10; Length 247;
 Best Local Similarity 72.6%; Pred. No. 1.2e-20;
 Matches 170; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Oy 1854 cctctagctctcttggggaaccacagagcctgctgctgctgctgctgctgctgct 1913
 Db 9 CCAACAGGAGATGACAGCGCTTCCCGCCACTGAGATTTCTGCTGCTGCTGCT 68
 Oy 1914 tgcctctgtttccacatctctgctgctgctgctgctgctgctgctgctgctgct 1973
 Db 69 TTTCTCGGGTTTGCATTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 128
 Oy 1974 catctcaagtttccacacaggtacatctctacacagctgctgctgctgctgctg 2033
 Db 129 CATCTTCAAGTTTCTTACACAGGTACCCACATTTCTACACACTTGGCCCAACATG 188
 Oy 2034 tgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2087
 Db 189 TGCCGGAATATTCGTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242

RESULT 15
 CNS02BUU 973 bp DNA linear GSS 12-MAY-2000
 LOCUS Tetradon nigriviridis genome survey sequence PUC-ori end of clone
 DEFINITION 254011 of library G from Tetradon nigriviridis, genomic survey
 sequence.
 ACCESSION AL190335.1 GI:7828439
 VERSION AL190335
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetradon nigriviridis.
 ORGANISM Tetradon nigriviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 973)
 Roest-Crollius, H., Jallion, O., Dasilva, C., Fizes, C., Fisher, C.,
 Bouneau, L., Billault, A., Queller, F., Saurin, W., Bernot, A. and
 Weissenbach, J.

TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetradon nigriviridis
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 973)
 Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fizes, C., Mincker, P., Brotier, P., Queller, F.,
 Saurin, W. and Weissenbach, J.

AUTHORS Human gene number estimate provided by genome wide analysis using
 Tetradon nigriviridis DNA sequence
 Unpublished
 3 (bases 1 to 973)
 Genoscope.
 JOURNAL Direct Submission
 REFERENCE Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases
 TITLE This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetradon nigriviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetradon.

FEATURES
 source
 1..973
 /organism="Tetradon nigriviridis"
 /db_xref="taxon:99883"
 /clone="254011"
 /clone_lib="G"
 /note="Genoscope sequence ID : C0AG254AH06SP1-end ;
 PUC-ori"

BASE COUNT 184 a 306 c 247 g 227 t 9 others

ORIGIN

Query Match 4.9%; Score 124.2; DB 12; Length 973;
Best Local Similarity 51.4%; Pred. No. 1.3e-18;
Matches 361; Conservative 4; Mismatches 322; Indels 16; Gaps 3;

```
QY 1626 agatgggacacgttggaagcagacgtgtccgcgcaactgtyggttttggcctt 1685
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 AGAGGGCTCTCAAGGAAAGGAACTCCAGCTGGCTCCAAACAACTGCTTCTCTG 338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1686 gcgtgagcacacacgtctgtggtgctgctgacgtacacagctgctgctgctgct 1745
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GAAGGACACCTTTGGGCGTTTCATGCGCTTTCACACCTGGGCACTCCTGTGCT 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1746 tggagctgctgctgctgttgcctgacacacacacacacacacacacacacacac 1805
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GTWGAACATGGCTGCTTCTGTACAAAGCGGACACTCGGTGGTGAAGGACAGGGGG 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1806 cgcctgctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1865
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 ACCCTTTAGCCCAAGCAAGCACTCTTCTCTCTGCTATCATCATCATCATCATCATCAT 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1866 ctltggggaaccacaaagcctgctgctgctgctgctgctgctgctgctgctgct 1925
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 TGTGGGCGAGCMACAGACATCCAGTGCAGAGCCCGCAGGTCCTCGGATCAGCTT 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1926 caccatctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1985
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 CACTCTGTGTATCTGTATCTGTATCTGTATCTGTATCTGTATCTGTATCTGTATCT 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1986 ttccacacaggttacatattctacacgctggttccaaacacaggtgctgctgtt 2045
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CAACCCCGGGCTGC-----AGGACGTGCTGAGAGATCTATTCGGCCGTACGCCAT 689
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2046 tgtatgatcagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2105
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 690 CATCACCGCTGTGCGTGTGCAAGAGTCATGTGCTGTGCTGTGCTGTGCTGTGCTGTG 749
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2106 cccacgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 750 CCCGTACATACACCATCGGGCAAGCAGCACCTCTTTGATG-----CACTGGCACGA 803
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2166 gaccacacacacacacacacacacacacacacacacacacacacacacacacacac 2225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 804 AAACCTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 863
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2226 -cttgcctgagctacacacacacacacacacacacacacacacacacacacacac 2284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 864 GCTTCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 923
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2285 tcaactttagcctgcttcaacttcaacttcaacttcaacttcaacttcaacttcaactt 2327
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 924 TCACCTTCAGCATGCTGCTGTACCTCATTTCTGCTGCTCTT 966
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: June 11, 2002, 05:49:17
Job time: 7093 sec

APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 488-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-484-565-1

Query Match 5.8%; Score 147.6; DB 1; Length 5275;
Best Local Similarity 47.5%; Pred. No. 1.7e-28;
Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;

Qy 1426 ctaaacataaataagacccaatccagtcgacggaagaacacaggctgctaagct 1485
Db 2078 CTCTTCATCATGATGAAAAATCTGTGAGTGTGATCTCTCAAGGAGGTGCTTTCGCC 2137
Qy 1486 gtgtgtccagcagcagtcgtctgaaggac---cagcgagtggttaagggttccatac 1542
Db 2138 AACTGACATCGACAGCTGCTGGCAGGACCAAGAAAGATCATTCAGGGGAGGCCACAC 2197
Qy 1543 tgcctgttgagtggtgctgctggtgagtcgacccctcccaagaagtgacctctac 1602
Db 2198 TGTCTCTTTGAGTGTGTGAATGTCTTATGAGGAGTACAGCGACGACAGATCAAGT 2257
Qy 1603 agatgccagcctgtggaagaagtgagcaccctgagggaggaagcagacctgtcccg 1662
Db 2258 GCGTGTGATGATGCGCCGATGAGACTTGTGTCATGAGAACACACTTCGTCGTCGCC 2317
Qy 1663 cgcacgtgtgtgttcttggtcttgagcagacacctcttggtgtgtgtgagctaac 1722
Db 2318 AAGGAGATCGAGTTCTCTGCGACCGACCTTCGGGATCGCACTGACGCTTCTTCT 2377
Qy 1723 agcgtgcgtgcgtgctgctggtgagcctggtgctgttgcctgacactaacac 1782
Db 2378 GTGCTGGGCAATTTCTTCAAGCCTTCGTCTGGGCGTCTTCATCAAGTTCGCAACAG 2437
Qy 1783 cctgtgtgagtcagcagggggccgctgtgcttcttctatgctggtgctccctgacga 1842
Db 2438 CCGATCGTCAAGGCGACCAACCGGAGCTCTCTATCTCTCTCTCTCTCTCTCTCTG 2497
Qy 1843 ggtgtagcagcctctatgcttcttggtggaaccacaaggcctgcgtgcttctaacg 1902
Db 2498 TGCTTCTCCAGCTCCCTCTTCTTCACTCGGGAGCCCGAGGATGAGCGCTGCGC 2557
Qy 1903 cagggccctcttgctggttcttccatcttccctgctcctgcctgacaggtgcctcaltc 1962
Db 2558 CAGCGGCTTTGGCATCAGCTCTCGTCTGTCATCTGTGCATCTGTGTAAGAAACAT 2617
Qy 1963 caactaacatcatcttcaagtttccacaaggaactaacctaacacagcctgagtc 2022
Db 2618 CGGTCTCTCTGCTGTGAG---GCCAAGATTCACACAGCTTCCACCGGAAGGAGG 2674
Qy 2023 caaaacacagtgctgagcctgttcttgatgacagtcagcagccagcctgctatctgt 2082
Db 2675 GGGGTCAACCTGACAGTTCCTGCTGCTTCTCTGACCTTCAATGACATGTCATCTGT 2734

Qy 2083 ctaactggcgtgtgtgtgtgaacccactgctgctgaaggaaataccagcgttccccc 2142
Db 2735 GCCATTGGCTCAATACAGACGCCCTCCCTGAGCTACCCGACACAGAGCTGGAGACAG 2794
Qy 2143 ctgtgatgtctgagtcagcagagcaactccctgggtcttactactgcttccctctac 2202
Db 2795 ATCATCTTCAATCACCCTGCGACAGAGGCTGCTCATGGGCGGCTTCTGATCGGCTAC 2854
Qy 2203 aatggcctctctccatgaagtcgttctgctgacgtacacgtggtgaagacttcagag 2262
Db 2855 ACCTGCTTGGCGCCGACATCGCTTCTTCCGCTTCAAGTCCCGGAGCTGCGACAG 2914
Qy 2263 aactacaagggccaagtgtgtaccttgaagcctgagccttcaactgtgttccgggtc 2322
Db 2915 AACTTCAATGAAGCCAAAGTTCACTACCTTCACATGCTCATCTTCTTCACTGTGATC 2974
Qy 2323 gctcttccacacagcagcgtctacagcaggaactgactgctgctgagcacaatgag 2382
Db 2975 TCTTTCATCCCGGCTTACGCGCAGCACTTACGCAAGTTCGTCTGCGGTGAGGTGATC 3034
Qy 2383 gctggcctgagcagcctgagcagcgtctggtgtgtatcttctgctcaagtgcagtg 2442
Db 3035 GCCATCTGGGCGCCAGCTTGGCTTGGCTTGGCTTATCTTCTTCAACAAAGTCTACATC 3094
Qy 2443 atccctgcgcgcagacctcaacagacagagcacttcc 2482
Db 3095 ATCTCTTCAAGCTTCCCGGACACCATCGAGAGGTGC 3134

RESULT 3
US-08-480-751-1
Sequence 1, Application US/08480751
Patent No. 5856684
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,751
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389

FILED DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-480-751-1

Query Match 5.8%; Score 147.6; DB 2; Length 5275;
Best Local Similarity 47.5%; Pred. No. 1.7e-28;
Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;

QY 1426 ctaaacataatgaagacaaatccagtgagcaaggaaggaacacagtgcttaagtc 1485
DB 2078 cttcttatcatatgataaaaaattctgtgagatgtctcaaggagagctcttctcc 2137
QY 1486 gtgtgtccagcgactgtcttgaaggacac---cagcgaagtgttcaagttccatca 1542
DB 2138 AACTGCAGTCGACACTGCGTGGGAGGACGAGAAAGATATCTGAGGGGAGCCACG 2197
QY 1543 tgcgtcttgagatgtgtccctgtgaggtgagccttcccaagaagtgaccttacc 1602
DB 2198 TGTCTGCTTTGAGTGTGTGAGATGCTCTGATGGGAGTACGACGACGATGCAAGT 2257
QY 1603 agtgcagagcgtgtggaagaagagtgagcactgaagggaagcagacctcttcccg 1662
DB 2258 GCGTGTGATAGTGCCTGTGACTTGTGTCATGATGAGAACACACACTTCTGATCGCC 2317
QY 1663 cgcactgtgtgttcttgcgttgcgtgagcagacctcttgggtgtcgtgagactaac 1722
DB 2318 AAGGAGATGAGATTCTGTGTCGAGACGACCTTCGGGATGACACGATCTTGTGCT 2377
QY 1723 agcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1782
DB 2378 GTGCTGTGGGATTTCTTCACAGCCTTCGTGCTGGGCTCTTCATCAAGTCCGACACG 2437
QY 1783 cctgtgtgtgagtcagcagggggcgccgtgtgttcttcttcttcttcttcttctt 1842
DB 2438 CCCATGTCAAGGCAACACCGGAGCTCTCTATCTCTCTTCTCTCTCTCTCTCTCTGC 2497
QY 1843 ggtagtgcagcctctatagctcttcttcttcttcttcttcttcttcttcttcttct 1902
DB 2498 TGTCTTCAGACTCCCTGTTCTTCATCGGGAGCCCGACAGACTGAGAGTCCGCTGCGC 2557
QY 1903 caggcctcttgccttggttccacatcttccctcgtcgtcgtcgtcgtcgtcgtcgt 1962
DB 2558 CAGCCCGGCTTTGGCATACAGCTTCTGCTGTGATCTCGGATCTCGTGAACCAAT 2617
QY 1963 caactatatacatcttcaagtttccacaaggtactctacatcttccacgcgtcgtgc 2022

DB 2618 CGGCTCCTCCTGCTGTTTGAG---GCCAAGATCCACACAGCTTCCACCGAAGTGTGG 2674
QY 2023 caaacacagcgtcgtgcctgttcttgaatagctcagcgcagcgtctatctgt 2082
DB 2675 GGGCTAACCTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2734
QY 2083 ctacttgcctgt 2142
DB 2735 GCCATTGGGTCAATACAGGCGCCCTCGAGCTACCGCAACACAGCTGAGAGAGAG 2794
QY 2143 ctgtgtatctgtatgtcagcagagagacacacacacacacacacacacacacacacac 2202
DB 2795 ATCATTTTATCATCCTGACGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2854
QY 2203 aatgcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2262
DB 2855 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2914
QY 2263 aactaacagcagcacaatgtgtcacttcaagcctcctcctcctcctcctcctcctcctc 2322
DB 2915 AACTCAATGAAGCAAGTTCATCAGCATCTCAGCATCTCCTTCTTCAATCGTGTGATC 2974
QY 2323 gcttcttcaacacagcagcgtctcagcagcagcagcagcagcagcagcagcagcagc 2382
DB 2975 TCTTTATCCCGCTACGACGACACTTACGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 3034
QY 2383 gctgagcgtgagcagcgtcagcagcagcagcagcagcagcagcagcagcagcagcagc 2442
DB 3035 GCCATCTGCGGCGGCGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 3094
QY 2443 atctctgcgcgcagcactcaacagcagcagcagcagcagcagcagcagcagcagc 2482
DB 3095 ATCTCTTCAAGCCTTCCGGAACACATCAGAGAGGTGC 3134

RESULT 4
US-08-943-986-1
; Sequence 1, Application US/08943986
; Patent No. 5962314
; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,986
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/484,565
; FILING DATE: 7-June-1995
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994

FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 209/069
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-353-784-1

Query Match 5.8%; Score 147.6; DB 3; Length 5275;
Best Local Similarity 47.5%; Pred. No. 1.7e-28;
Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;

QY 1426 ctaaaataatgagacacaaatccagtgcaaggaagacaccagtgctctaagtc 1485
DB 2078 CTCCTTCAATGATGATAAAATTCGTGAGTTCATCAAGGAGGAGGCTTTCTCC 2137
QY 1486 gtagtgcagcagctgtctgaagggcac---cagcagatggttacgggtttccatcac 1542
DB 2138 AACTGACAGTCGAGACTGCTGCGAGGAGACAGAAAGGATCTTGAAGGGGAGGCCACC 2197
QY 1543 tgcgtctgaagtgtgtccctgtgaggtcgtggaaccttcctcaacaagagtgaacctctac 1602
DB 2198 TGCTGCTTGAAGTGTGAGATGCTCTGATGGGAGTACAGCAGACAGATGCAAGT 2257
QY 1603 agatgcagcctgtggaagaagagtggcaactggaaggaagcaagacctgtctccg 1662
DB 2258 GCGTGTGATGAAGGCCCTGATGCTTGTGTCATGAGAACACACTTCTGCATCGCC 2317
QY 1663 cgcacgt 1722
DB 2318 AAGGAATGAGATGATTTCTGTCGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2377
QY 1723 acgt 1782
DB 2378 GTGCTGGGATTTTCTTCAACAGCCTCGTGTGGGCTGTCTATCAAGTTCCGCAACAGC 2437
QY 1783 cctgt 1842
DB 2438 CCATGATGATGAAGCCCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2497
QY 1843 gtagtgcagcctgt 1902
DB 2498 TGCTTCTCAAGCTCCCTGTTCTTCAATCGGAGGAGGCCAGAGACTGAGCTGCCTGCGCG 2557

QY 1903 cagggcctcttggccctgt 1962
DB 2558 CAGCGCGCTTTGGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2617
QY 1963 caactaatcatcatctcaagtttccaccagaagtaactcatcttaccagcctgtgtgtgtgtgt 2022
DB 2618 CGGCTCTCTCTGCTGT 2674
QY 2023 caaaacacggt 2082
DB 2675 GGGCTAACCTGACATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2734
QY 2083 ctaactgt 2142
DB 2735 GCATTTGGCTCAATACAGCGGCCCTCGAGCTACCGAACAGAGCTGAGAGCAGAG 2794
QY 2143 ctagtgaatgt 2202
DB 2795 ATCATCTTCAATCACTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2854
QY 2203 aatgacctctctccatcagtgctgt 2262
DB 2855 AACTGCTTGTGCTGGCGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2914
QY 2263 aactaacagagcccaatggt 2322
DB 2915 AACTCAATGAAGCAAGTTCAATCAGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2974
QY 2323 gctcttaccacagcagcagcgt 2382
DB 2975 TCTTTCATCCCGCCCTACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 3034
QY 2383 gctgagctgt 2442
DB 3035 GGCATCTGTGGCGGCGAGCTTTGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3094
QY 2443 atctctgcgcgcagcctcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2482
DB 3095 ATCCTCTTCAAGCCTTCCCGGAACACCATGAGAGAGTGC 3134

RESULT 6
US-08-484-719B-1
; Sequence 1, Application US/08484719B
; Patent No. 6031003
; GENERAL INFORMATION:
; APPLICANT: Edward F. Nemeth, Edward M.
; APPLICANT: Brown, Steven C. Hebert,
; APPLICANT: Bradford C. Van Wageningen,
; APPLICANT: Manuel F. Balandrin,
; APPLICANT: Forrest H. Fuller, Eric G.
; APPLICANT: Delmar, Scott T. Moe
; TITLE OF INVENTION: CALCIUM RECEPTOR ACTIVE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS Word
; SOFTWARE: FASTSEQ for Windows Version 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,719B
; FILING DATE: 7 June, 1995
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Douglas C. Murdock
REGISTRATION NUMBER: 37,549
REFERENCE/DOCKET NUMBER: 213/007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
US-08-484-719B-1

Query Match 5.8%; Score 147.6; DB 3; Length 5275;
Best Local Similarity 47.5%; Pred. No. 1.7e-28;
Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;

Qy 1426 ctaaacataatgagacccaataatccagctgagcaggaagaacaccaggtgcctaagctc 1485
Db 2078 CTCTTCAATCAATGAGAAATTTCTGTGAGCTGATTCTCAAGGAGGTGCTTTCTCC 2137
Qy 1486 gtgtgtccagcagactgtcttgaaggacac--cagcgagtgttaacgggttccatcac 1542
Db 2138 AACTGCATTCGAGACTGCTGCGACGAGGACGAGAAAGCAATTCGAGGGGAGGCCACG 2197
Qy 1543 tgcgtcttgagtggtgcccctgagggcgtggagaccttccctcaacaagagtgaaccttac 1602
Db 2198 TGCTGCTTTGAGTGTGTGAATGTCTGATGGGAGTACAGCGACGACAGATCAAGT 2257
Qy 1603 agatgccgcgcttgaggaaagaagatgggagaccttgaggaggaagcagaccgtctccg 1662
Db 2258 GCCTGTGTAAGTGCCTGATGACTTGTGTCGAATGAGAACCACTTCCTGCAATCGGC 2317
Qy 1663 cgcacatgtgtgttttcttgcttgcttgacacacctcttggtgtgtctgctgacgtaac 1722
Db 2318 AAGGAGATCGAGTCTTGTCTGCGACGAGCCCTTCGGATCGCACTACGCTTTTGT 2377
Qy 1723 acgctgtgctgctgctgctgctgctgagactgctgctgttgcctgagcctagaacacc 1782
Db 2378 GTGCTGGGCAATTTCTCAACAGCCTTCGTGCTGGCGCTTCATCAAGTTCCGCAACAG 2437
Qy 1783 cctgtgtgtgtgtgagcagggggcgccgtgtgtcttctatgtctgtggcctctgacaga 1842
Db 2438 CCGATCTCAAGGCGCAACCGGAGGCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTGC 2497
Qy 1843 ggtagtggcagcctctatgtcttcttgaggagaccacaagcctgctgtctgtacgc 1902

Db 2498 TCTTCTCCAGCTCCCTGTTCTTCAATCGGGAGGCCACGAGACTGACGTCGGCCTGCGC 2557
Qy 1903 cagggcccttggccctgtgttcaacatctctcgtctgctgacagttgcctcatc 1962
Db 2558 CAGCGCGCTTGGGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2617
Qy 1963 caactaatcatatcttcaagtttccaccaggaagtaacctcaacttcaacagcctgggac 2022
Db 2618 CCGGCTCTCTGCTGTTTGTAG--GCCAAGATTCCACCACTTCCACCGGAAGTGTGG 2674
Qy 2023 caaacaccagtgctgagcctgttctgatatcagctcagcagcagcagctgtatctgt 2082
Db 2675 GGGCTCAACCTGACGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2734
Qy 2083 ctaactgtgctgtgtgtgtgagcccaactgctgctgaggaataccagcgtctcccat 2142
Db 2735 GCCATTTGGCTCAATACAGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2794
Qy 2143 ctgtgtatgtctgtgtgtgacagagcaactctccctgggtctatctgctcctctac 2202
Db 2795 ATCATCTTATACCTGCGCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2854
Qy 2203 aatggcctcctccatcagtgctgtgctgtgagctgtgtaagactgtgcagag 2262
Db 2855 ACCTGCTTCTGCGCGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2914
Qy 2263 aactacaagagagcaaatgyltcaacttcaagctgtcttcaacttctgtctgtgac 2322
Db 2915 AACTTCAATGAGCAAGTTCACTCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2974
Qy 2323 gcttcttaccagagcagcgtctacagcaggaagtaactgtcctgtgcgcacatgag 2382
Db 2975 TCTTTCATCCCGCCCTACCGCACCTTACGCAAGTTGCTGCTGCTGCTGCTGCTGCTG 3034
Qy 2383 gctggctgagcagcctgtgagcagcgtctggtgtgtatcttctgctaagtacgtg 2442
Db 3035 GCCATCCTGCGCGCGCGCGCTTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3094
Qy 2443 atcctctgcgcagcagacctcaacagcagcagcagcagcagcagcagcagcagcagc 2482
Db 3095 ATCTCTTCAAGCTTCCCGGAACACCATCGAGGAGGTGC 3134

RESULT 7
US-08-484-159-1
Sequence 1, Application US/08484159
Patent No. 6313146
GENERAL INFORMATION:
APPLICANT: Bradford C. Van Magenen
APPLICANT: Manuel F. Balandrin
APPLICANT: Eric G. Del Mar
APPLICANT: Edward F. Nemech
TITLE OF INVENTION: CALCIN RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,159
FILING DATE: 7 June, 1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 214/101
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-484-159-1

Query Match 5.8%; Score 147.6; DB 4; Length 5275;
Best Local Similarity 47.5%; Pred. No. 1.7e-28;
Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;

QY 1426 ctataacataatgagaccaaataccagtgacgagaaaggaacacaggtgcttaagct 1485
DB 2078 ctcttcatcatgataaataatcttgatgagatgattctcaaggagctcttctcc 2137
QY 1486 ggtgttcagcagcagctcttgaaggacac---cagcagtggttaccggtttccatcac 1542
DB 2138 AACTGAGTCGACGACCTGCGGAGGAGACGAAAGGAATCAATTAGGGGAGCCACAC 2197
QY 1543 tgcgtcttgagtggtgctccttgaggcttgagaccttcccaacaagaatgaccttacc 1602
DB 2198 TGTGTGTTGAGTGAGTGATGCTGATGGGAGTACAGCCAGACGACAGATGCAAGT 2257
QY 1603 agatgcagcctctgtgagaaagaagtgaggacactgagggaaagcagacgtcttccg 1662
DB 2258 GCCTTGATGAAGTGGCCCTGACTCTGTGTCATGAGAACCACTTCCTGATCGCC 2317
QY 1663 cgcacgtgtgttttggcttgagcacacactcttggtgtgtctgagcagctaac 1722
DB 2318 AAGGAGATGAGATTGTGCTGAGACGACCCCTTGCGGATCGACTACACGCTTTGCT 2377
QY 1723 acgctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1782
DB 2378 GTGCTGGGATTTTCTCTACAGCCTTCGTCTGGGCTCTTCAATCAAGTTCGCCAACG 2437
QY 1783 cctgtgtgagtgacagcagggcgagcgtgtgttcttatgtctgtcgtcgtcgtcgtcgtc 1842

DB 2438 CCCATCTCAAGGCGACCAACCGGAGCTCTCTATCTCTCTCTCTCTCTCTCTCTCT 2497
QY 1843 gtagtgagcagcctctatgcttcttggggaaccacaagcctgtgtcgtcagc 1902
DB 2498 TGCTTCTCCAGCTCCCTGTTCTTCATGCGGAGCCCGAGACGACGCTCCCTCTCC 2557
QY 1903 cagcgcctcttgccttggtttccacatcttcctgcctcgtcgtcgtcgtcgtcgtc 1962
DB 2558 CAGCGGCTTTTGATGATGACCTTCGTCTGATCTGCTGATCTGCTGATCTGCTGAT 2617
QY 1963 caactatctatcttcaagtttccacaaggtactactatcttccacagcgtggtc 2022
DB 2618 CGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2674
QY 2023 caaacacagtgctgagcgttgtgtatgatcagctcagcagcagcagcagcagcagc 2082
DB 2675 GGGCTCAACCTGACGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2734
QY 2083 ctactgt 2142
DB 2735 GCCATTTGGCTCAATACAGCGCCCTCGAGCTACCGCAACGACGAGCTGAGAGAG 2794
QY 2143 ctgt 2202
DB 2795 ATCATCTTCATCACTGCGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2854
QY 2203 aatgctcctctccatcagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2262
DB 2855 ACCTGCTGCTGCGCGGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2914
QY 2263 aactacaagagccaaatggtacaccttcagcgtcgtcgtcgtcgtcgtcgtcgtcgt 2322
DB 2915 AACTCAATGAAGCAAGTTCATCCTTCAGATGCTCATCTTCTTCATCGTGTGATC 2974
QY 2323 gcttcttcaacagcagcagcgtctcagaaggaagtaactgctcgtcgtcgtcgtcgt 2382
DB 2975 TCTTTATCCCGCCCTACGACGACCTTACGCAAGTCTGCTGCTGCTGCTGCTGCT 3034
QY 2383 gctggtctgagcagcctgagcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2442
DB 3035 GCCATCTGCGGCGGACCTTGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3094
QY 2443 atcctctgcccagacactcaacagcagcagcagcagcagcagcagcagcagcagc 2482
DB 3095 ATCCTCTTCAAGCCTTCCGGAACACCATGAGAGGTGC 3134

RESULT 8
US-08-485-588-3
Sequence 3, Application US/08485588
Patent No. 5688938
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,588
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-485-588-3

Query Match          5.6%; Score 141.2; DB 1; Length 3809;
Best Local Similarity 47.1%; Pred. No. 6.4e-27;
Matches 499; Conservative 1; Mismatches 554; Indels 6; Gaps 2;
```

```
Db 2233 GTGCTGGGCAATTTTCTCTACAGACGCTTTGCTGGGTGATTTATCAATTCGGACACACA 2292
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 1783 cctgtgtgagtgtaagcaaggggagccgctgtgtcttcttattatgtgtgtccctgcagca 1842
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2293 CCCATTGTCAAGGGCCACACACGAGAGCTTCCTACCTTCCTTCCTTCCTTCCTTCCTTCCT 2352
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 1843 ggtatggagcaccctctataggtcttcttggggaacccacaagcccggtgtgttgaagc 1902
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2353 TGCCTTCCTCAGCTCCCTTTCTTCTCATCGGGAGGCCCCAGAGACTGAGAGCTGCGCTGCGC 2412
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 1903 caggccctcttcttcccttcttcttccatcttcccttcttcttcccttcttcccttcttcc 1962
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2413 CAGCGCGGCTTTGGGATCAGCTGTGCTGTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 2472
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 1963 caactaatcatcatcttcaagtlttccacaaagtaacctaatcttaccacagcttggctc 2022
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2473 CGTGCTCTCTCGGTGTGTGAGG--CCAAAGATCCCCACACAGCTTCACCGCAAGGTGTGG 2529
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 2023 caaaacacagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2082
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2530 GGGCTCAACCTGCAAGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2589
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 2083 ctaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2142
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2590 GTGATGTGCTCTACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2649
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 2143 ctgtgtatgtctttagtgacagagagacaactccctgtgtgtgtgtgtgtgtgtgtgtgt 2202
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2650 ATCATCTTCATCAGCTGTGCAAGGGCTCCCTCAAGGGCTCCGAGGCTTCCTGAGGCTCAG 2709
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 2203 aatggcctcctctccatcagtgagcttctgtcagctcctgtgtgtgtgtgtgtgtgtgtgtgt 2262
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2710 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2769
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 2263 aactcaacagagcccaaatgtgtacacttcaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2322
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2770 AACTTCATGAAGCCAAAGTATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2829
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 2323 gcttcttcaacagagccagtgcttcaagagagagagagagagagagagagagagagagagatg 2382
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2830 TCCCTTCATTCAGGCTGTATGCCAGCACCTATGCAAGTTGTCTGCTGCTGCTGCTGCTGCTGCT 2889
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 2383 gctggagctgagcagcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2442
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2890 GCCATCTGCGAGCCAGCTTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2949
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 2443 atcctctgcgcgcccagacctcaacagcacagagacacttc 2482
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2950 ATTCTCTCAAGCCATCCCGACACACATCGAGGAGGTGC 2989
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 9
US-08-484-565-3
: Sequence 3, Application US/08484565
: Patent No. 5763569
: GENERAL INFORMATION:
: APPLICANT: Edward M. Brown
: APPLICANT: Steven C. Hebert
: APPLICANT: James E. Garrett, Jr.
: TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
: TITLE OF INVENTION: MOLECULES
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: First Interstate World Center
: STREET: Suite 4700
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
```

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,565
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-484-565-3

Query Match	5.6%	Score 141.2	DB 1.2	Length 3809
Best Local Similarity	47.1%	Pred. No. 6.4e-27		
Matches	499	Conservative 1	Mismatches 534	Indels 6
			Gaps 2	
QY 1426	ctaaacataaatgagaccanaatccagctgagcaggaagaaacacagctgcttaagct	1485		
Db 1933	CTCTTCATCAACGAGAGAAATCCTGTGAGATGGGTCTCCAGGAGAGTGCCCTTCTCC	1992		
QY 1486	gtgtgttccagcagctgtctcttgaaagagcaccagcag---tgtttaagtggtttccatcac	1542		
Db 1993	AACTGCACCCGAGATGGCTGTGCACAGGACCAAGAAAGGATCATTTGAGGGGAGGCCACAC	2052		
QY 1543	tgctgcttggatgtgtgtccctgtgggttgagacttccccaagaagtgtaacctcac	1602		
Db 2053	TGCTCTTTGAGTGTGTGGAGTGTCTCTGATGGGGAGTATATGATGAGACAGATGCCACT	2112		
QY 1603	agatgcccagccttgtggaagaagaagtgaggacctgaggaagcacaagcctgtctccgg	1662		
Db 2113	GGCTGTAAACAAGTCCCAAGATGACTTCTGTGGTCCATTTAGAACACACACCCTCGCATTTGCC	2172		
QY 1663	cgcactgtgtgttttcttgcttgcgttagcacaacctcttggtgtctgctgtgcagactaac	1722		

[illegible]

US-08-480-751-3
Sequence 3, Application US/08480751
Patent No. 5856684
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth
APPLICANT: Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street

CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,751
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-480-751-3

Query Match	5.6%	Score 141.2	DB 2	Length 3809
Best Local Similarity	47.1%	Pred. No. 6.4e-27		
Matches 499	Conservative 1	Mismatches 554	Indels 6	Gaps 2
QY 1426	ctaacaataatgagaccacaaatccatgagtcacggaagaagacaaccagtgctcaatct	1485		
Db 1933	CTCTCATCAACAGAGAGAAAATCCTGTGGATGGGTTCTCCAGGAGAGTGCCCTTCTCC	1992		
QY 1486	gtgtgtgtccagcagactgtctctgaagggcaccagcagag--tggttaagggtttccatcac	1542		
Db 1993	AACGCAACCCAGACTGGCTCTGGCAGGACCAAGAAAGGATATTGAGGGGAGGCCACAC	2052		
QY 1543	tcgtgccttgaagtgtgtccctgtgaggcttggaaccttcctcaacacagatgactctac	1602		
Db 2053	TGCGTCTTTAGATGTGTGTGGAGTGTCTCATGTGGGAGTATATGATGATGACAGATGCCACT	2112		
QY 1603	agatgcacagccttgtgggaaagaagatgaggcaaccttgagggaagccagacctgtctcccg	1662		

Db	2113	GCCTGTAAACAAGATGCCAGATGACTTCTGGTCCATATGAACAACCAACTCTGCACATTGCC	2172
QY	1663	cgcactgtggatgttttttggatcttgcgttggacaacctcttgggtgtctgtgcagactaac	1722
Db	2173	AAGAGATACGAGTTTCTGTGTGGAGAGAGCCCTTTGGGATGGCACTACACCTCTTTGCC	2232
QY	1723	aagctgtcgtcgtcgtcgtcgtcgttggaaactcgtcgtgttctgcctcgtgaactagaac	1782
Db	2233	GTGCTGGGCAATTTTCCCTGACAGGCTTTGAGCTGGGTGTATATCAAGTTCCGGCAACCA	2292
QY	1783	ccctgtgtgaggtcagaagaggcgccgctgigtcttctatgcttgagtggtctccctgacga	1842
Db	2293	CCCAATTGCAAGGCGACCAACCAAGAGCTCTCTACCTCTCTCTCTCTCCCTGCTTGC	2352
QY	1843	ggtagtgtagcagccctctatgtgccttcttcttggaaaccacaaagccctgcgttgccttgc	1902
Db	2353	TGCTTTCCAGACTCCCGTCTTCTCATGCGGAGACCCCAAGACTGGAGCTGGCGGCTGGCG	2412
QY	1903	cagagccctcttgcctctgtgttcacacatctctcgtcctgcctgcagagttcgctcatlc	1962
Db	2413	CAGCGGCGTTTGTGGCATCACTCTTGAGCTCTGATCTCATGTGATCTCTGTGTGAAAACAC	2472
QY	1963	caactatcatcatcattcacaagttctccacaaggtacacatctacacagccgtgggtc	2022
Db	2473	CGTGTCTCTCGGTGTTTGG---CCAGATCCCAACAGCTTTCACCGCAAGTGGTGG	2529
QY	2023	caaaacaaggtgtgcgtgtgttggatgatcagtcagcgccacagctgcttactgt	2082
Db	2530	GGGCTCAACCTGCAAGTCTCTGTGTGTTTCCCTGCACCTCATGCAAGTGTCAATCTGT	2589
QY	2083	ctaactgtcgtgtgtgtgtgtgtagcccaactcgtcgtcgtcagaagataccagtgctccccat	2142
Db	2590	GTGATCTGGGCTCTACACCGGCGCCCTCCCTCAAGCTACCGAACAGAGCTGGAGGATAG	2649
QY	2143	ctgtgtgtagtgcgt	2202
Db	2650	ATCATCTTTCATCAAGTGTCCACAGAGGCTCCCTCATGTGGCCCTGGCTTCTCATGGCTAC	2709
QY	2203	aatgtcctcctctcatcagtgctccttgcctgcagctacatcgtgtgaagtaactgtcagag	2262
Db	2710	ACCTGCTGCTGGCTGTGCATCT	2769
QY	2263	aactacaagagagcaaatgtgtcaactcttcaactcagctgtcttcaacttgttctcgtgtc	2322
Db	2770	AACCTTCATGAAGCAAGTGTCTACCTTCATCGAGTATCTTCTTATCTCTGTGATTC	2829
QY	2323	gctcttctcaacacagcgagcgtctcagaagcaagtactcgtcgtcgagcaacatgtgt	2382
Db	2830	TCTCTTCATTCAGCTATGCGACGACCTATGGCAAGTTGTCTCTGCGGTAGAGGTGATT	2889
QY	2383	gctgtgctgagacagctgtgagacagcgtgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2442
Db	2890	GCATTCCTGGGAGGACCACTTTGGCTGTGGGCGGTGCATCTTCTTCACAAAGATCTACATC	2949
QY	2443	atctctgtcgcccaagacctctcaacagcaagagcaacttc	2482
Db	2950	ATTCTCTTCAAGGCATCCGCCAACACCATTCAGAGAGAGTGC	2989

```

1 RESULT
2 US-08-943-986-3
3 ; Sequence 3, Application US/08943986
4 ; Patent No. 5962314
5 ;
6 ; GENERAL INFORMATION:
7 ;
8 ; APPLICANT: Edward M. Brown
9 ;
10 ; APPLICANT: Steven C. Hebert
11 ;
12 ; TITLE OF INVENTION: CALCULUM RECEPTOR-ACTIVE
13 ; TITLE OF INVENTION: MOLECULES
14 ; NUMBER OF SEQUENCES: 20
15 ;
16 ; CORRESPONDENCE ADDRESS:
17 ;
18 ; ADDRESSEE: Lyon & Lyon
19 ;
20 ;
21 ;
22 ;
23 ;
24 ;
25 ;
26 ;
27 ;
28 ;
29 ;
30 ;
31 ;
32 ;
33 ;
34 ;
35 ;
36 ;
37 ;
38 ;
39 ;
40 ;
41 ;
42 ;
43 ;
44 ;
45 ;
46 ;
47 ;
48 ;
49 ;
50 ;
51 ;
52 ;
53 ;
54 ;
55 ;
56 ;
57 ;
58 ;
59 ;
60 ;
61 ;
62 ;
63 ;
64 ;
65 ;
66 ;
67 ;
68 ;
69 ;
70 ;
71 ;
72 ;
73 ;
74 ;
75 ;
76 ;
77 ;
78 ;
79 ;
80 ;
81 ;
82 ;
83 ;
84 ;
85 ;
86 ;
87 ;
88 ;
89 ;
90 ;
91 ;
92 ;
93 ;
94 ;
95 ;
96 ;
97 ;
98 ;
99 ;
100 ;

```


APPLICANT: Bradford C. Van Meegenen,
 APPLICANT: Manuel F. Balandrin, Jr.
 APPLICANT: Forrest H. Fuller, Eric G.
 APPLICANT: Delmar, Scott T. Moe
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 TITLE OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: MS Word
 SOFTWARE: FASTSEQ for Windows Version 3.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,719B
 FILING DATE: 7 June, 1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/353,784
 FILING DATE: 9 December, 1994
 APPLICATION NUMBER: PCT/US/94/12117
 FILING DATE: 21 October, 1994
 APPLICATION NUMBER: U.S. 08/292,827
 FILING DATE: 23 August, 1994
 APPLICATION NUMBER: U.S. 08/141,248
 FILING DATE: 22 October, 1993
 APPLICATION NUMBER: U.S. 08/009,389
 FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/993,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Douglas C. Murock
 REGISTRATION NUMBER: 37,549
 REFERENCE/DOCKET NUMBER: 213/007
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELER: 67-3510
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3809 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 373..3606
 US-08-484-719B-3

Query Match	5.6%	Score 141.2	DB 3	Length 3809
Best Local Similarity	47.18	Pred. No. 6.4e-27		
Matches 499	Conservative 1	Mismatches 554	Indels 6	Gaps 2

QY 1426 ctaaacataaatgagaccataatccagtggcagcgaaagacaaccagtgccctaagtct 1485

Db 1933 ctttttcacacacgagagaaaatcctgtgagtggttctccaggagaggtgccctttccc 1992

QY	1486	gtatgttccaaagcagctctcttgaagagcaacaaagcag---tggtaagagttccatcac	1542
Db	1993	AACCTCAGCCAGCACTCCTCTGGCAGGGACCAGAAAGGATATTGAGGGGAGCCACC	2032
QY	1543	tgctgctttgagtgatgctgcctctgttgagtcgtggaccttccataaagatgacctac	1602
Db	2053	TGCTGCTTTGTGATGTGGAGGTGCTCCGATGAGGGAGTATGATGATGACACATGCCAGT	2112
QY	1603	agatgcaagcctgttgaggaaagaagatgggacccctgaaggaagcagacgtgctccg	1662
Db	2113	GCCTGTAAACAAGTCCCAAGATGACTTGTGGTCCAAATGAACAACACACCTCTCGCATTTGCC	2172
QY	1663	cgcactgtagtgatttttgatcttgcgttgagcaacactctctggatgctgtagcagctaac	1722
Db	2173	AAGGAGATCGAGTTTCTGTGTGGAGGAGGCCCTTTGGGATGCACTACCCCTTTTGCC	2232
QY	1723	acgctgctgctgctgctgctgctgttggaactgctgctgctgttcttgcttgcactagaacc	1782
Db	2233	GTGCTGGGCAATTTTCCGTGACAGCCTTTGTCGTGGGTGTTTATGAAGTCCGCAACCA	2292
QY	1783	ccctgtgtagtgatgaagaggagggccgcgtgtgtcttcttatgctggcctctggcagca	1842
Db	2293	CCCATTTGTCAAGGCCACCAACGAGAGCTCTACTCTCTCTCTCTCTCTCTCTCTGCTGC	2352
QY	1843	gtagtgagcagcctctatgctctcttcttgggaaaccacaagcgtgctgcttgtagcgc	1902
Db	2353	TGCTTTTCCACACTCCCTGTTCTTTCATTCGAGGAGCCCAAGATGAGACGTCGCGCTGCGC	2412
QY	1903	cagcgctcttttgcccttggttttcaacatcttctctgctgcctgaagttcgcctalc	1962
Db	2413	CAGCGGCGCTTTGGCATACCTTCTGCTCTGATCTCATGATCTCGTGTGAANAACAC	2472
QY	1963	caactaatcatcatctcaagattctcaccaagaatcctacatctacacagcctgggct	2022
Db	2473	CGTGTCCTCGTGATTTTGAAG--CCAAAGATCCCAACAGGTTCCACCGCAAGTGGTGG	2529
QY	2023	caaaacacaggtgctgctgctgtttgtgatatgctcagtcagcgagccagctgctatcgt	2082
Db	2530	GGGCTCAACCGCAAGTCTCGCTGTTTCTTCATGCACTTCATGCAAGTTTCTCATCTGT	2589
QY	2083	ctaaacttgctgctgctgctgctgagcccaactgctgctaggaataacagagcttcccac	2142
Db	2580	GTGATCTGGCTCTACACCGGCGCCCTCTCAAGCTACCGCAACAGAGAGCTGAGAGATAG	2649
QY	2143	ctggtgtagtcttgatgacagagaaacaaactccctgggctcatactgagcttccctac	2202
Db	2650	ATCATCTTCAATCAAGTGTCCACGAGAGGCTCCCTCATGGCCCTGGGCTTCTGTATCGGCTAC	2709
QY	2203	aatggcctccctctcatcagtgcttgccttgctgagtgactcctggtaaggaactggcagag	2262
Db	2710	ACCTGCTCGCTGGGTGCCATCTGCTTCTTTCCTTAAAGTCCCGGAAGCTGCGGAG	2769
QY	2263	aactacaacagcgcaaatgctgacacttcaagcctcttcaactctgctgctgagtc	2322
Db	2770	AACCTTCAATGAAGCAAGTTCATCATTACCTTGAGATGCTCATCTTCTTATCGCTCGATC	2829
QY	2323	gacttcttcaacaacgagccagcgtctacagcagcagaagtacactgctgctggcgcacaatg	2382
Db	2830	TCTCTTCAATTCAGGCTATGACAGACACTATGCGAAGTTTGTCTGCGGTAGAGGTGAT	2889
QY	2383	gctgtagtgtagcagcctcgagcagcagccttctggtggtatcttctgcctaagtgtcactg	2442
Db	2890	GCCATTCCTGGAGCGACACTTTGGCTTGTGGCTGTGATCTTCTTCAACAAAGATCTTATC	2949
QY	2443	atcctctgcgcgcagagcctcaacagcaagcaagcacttc	2482
Db	2950	ATTCTCTTCAAGCATCCCGCAACCACTTCAGAGAGAGTGC	2989

RESULT 14
US-08-546-998-2
; Sequence 2, Application US/08546998
; Patent No. 6211244


```

: STREET: Suite 4700
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: FASTSEQ
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,159
: FILING DATE: 7 June, 1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: PRIOR APPLICATION DATA: including application
: PRIOR APPLICATION DATA: described below: 9
: APPLICATION NUMBER: 08/353,784
: FILING DATE: 9 December, 1994
: APPLICATION NUMBER: PCT/US/94/12117
: FILING DATE: 21 October, 1994
: APPLICATION NUMBER: U.S. 08/292,827
: FILING DATE: 23 August, 1994
: APPLICATION NUMBER: U.S. 08/141,248
: FILING DATE: 22 October, 1993
: APPLICATION NUMBER: U.S. 08/009,389
: FILING DATE: 23 February, 1993
: APPLICATION NUMBER: U.S. 08/017,127
: FILING DATE: 12 February, 1993
: APPLICATION NUMBER: U.S. 07/934,161
: FILING DATE: 21 August, 1992
: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Heber, Sheldon O.
: REGISTRATION NUMBER: 38,179
: REFERENCE/DOCKET NUMBER: 214/101
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3809 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 373..3606
: OTHER INFORMATION:
: US-08-484-159-3

```

```

Query Match          5.6%; Score 141.2; DB 4; Length 3809;
Best Local Similarity 47.1%; Pred. No. 6.4e-27;
Matches 499; Conservative 1; Mismatches 554; Indels 6; Gaps 2;

```

```

QY 1426 ctaacataatgagaccaaataccagtgcaacggaagaacacagtgctcaagctc 1485
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1933 CTCTTCAATCAACGAGAGAAATCCGTGAGTGGCTTCTCCAGGAGGCGCTTCTCC
QY 1486 gtgtttccagagcagtgcttgaagaagcagcagcgag---tggttaagggttcacac 1542
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1993 AACTGACGCGGAGACTGCTGGCAGGAGACGAGAAAGGATCATGAGGAGGCCACC 2052
QY 1543 tgcgtcttgaagtggtgctcctggtgagacctctcaacaagaagtgacctacac 1602
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2053 TCGTCTTGAAGTGTGAGTGTCTCTGATGGGAGATATAGTATGACAGACATGCCACT 2112

```

```

QY 1603 agatgccagccttgtyggaagaagaagtgaggacctgaggaagcagacctgtcccg 1662
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2113 GCCTGTAAACAAGTGCCAGATGACTTCTGTGCCAATGAGAACCAACCTCTCATTTGCC 2172
QY 1663 cgcacgtggtgttttggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1722
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2173 AAGGAGATCAGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2232
QY 1723 acgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1782
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2233 GTGCTGGGAGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2292
QY 1783 cctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1842
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2293 CCCATTGTCAAGGCCAACCAACGAGAGCTCTCTTACTCTCTCTCTCTCTCTCTCTCTCT 2352
QY 1843 ggtagtgagagcctctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1902
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2353 TGTCTTCCAGCTCCCTGTCTTCTTCAATCGGGAGCCCGAGACTGAGCGTGGCCCTGCGC 2412
QY 1903 caggccctcttgcccttggtttacacatcttctgtgtgtgtgtgtgtgtgtgtgtgtgt 1962
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2413 CACCGCGGCTTTGGCATCAAGCTTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2472
QY 1963 caactaatcatcttcaagtttccaccaaggtactcaatcttcaaccaagcctgtgtgt 2022
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2473 CGTGTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2529
QY 2023 caaacaacagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2082
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2530 GGGCTCAACCTGCAATGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2589
QY 2083 ctaacttggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2142
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2590 GTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2649
QY 2143 ctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2202
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2650 ATCATCTTCAATCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2709
QY 2203 aatgscctctctcaatcagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2262
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2710 AACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2769
QY 2263 aactacaagagccaaatggttcaacttcaacttcaacttcaacttcaacttcaacttcaact 2322
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2770 AACTTCAATGAGCCAAAGTTCAATCACTTCAAGATGCTCATCTTCTTCAATGCTGTGATC 2829
QY 2323 gaccttcaaccaagcagcggtctagaagcgaagtactcgtcgtcgtcgtcgtcgtcgtcgtcgt 2382
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2830 TCCTTCAATTCAGCCATATGCAACGACCTATGCAATGCTTGTCTCTCCCTAGAGGTGATT 2889
QY 2383 gctggtgtgagagccttgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2442
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2890 GCCATCTTGACCAAGCTTTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2949
QY 2443 atctctgcgcaccagactcaacagcagcagcagcagcagcagcagcagcagcagcagcagc 2482
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2950 ATTCTTCAAGCAATCCGCAACACCATCAAGAGAGGTGC 2989

```

```

Search completed: June 11, 2002, 06:49:19
Job time: 6410 sec

```

1	2524.8	100.0	2526	22	AA0159501
2	2524.8	100.0	2951	22	AA0159505
3	2500.2	99.0	2526	22	AA0175116
4	2290.8	90.7	2292	22	AA0159502
5	2274.2	90.0	2333	21	AA550745
6	1889.4	66.9	2579	21	AA550744
7	1674.4	66.3	2771	21	AA550743
8	1001.2	39.6	1101	22	AA015903
9	908.6	36.0	8191	22	AA017515

XX

XX

04-OCT-2001.

28-MAR-2001; 2001WO-US09396.
28-MAR-2000; 2000US-192978P.
(LEXI-) LEXICON GENETICS INC.
Malke DM, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
Zambrowicz B, Sands AT;
WPI: 2001-616474/71.
P-PSDB: AAE11969.
Novel isolated polynucleotides encoding human G-protein coupled
receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
expressed NGPCRs for diagnosis of disease, and as probes or primers -
Claim 1; Page 71-72; 80pp; English.
The present sequence is a cDNA encoding human novel G-protein coupled
receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span
the cellular membrane and are involved in signal transduction after
ligand binding. The NGPCR polynucleotide sequences, are useful
in diagnosis and treatment of a disease involving NGPCR, for detecting
mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
disease, for screening drugs effective in treatment of symptomatic or
phenotypic manifestations of perturbing the normal function of NGPCR in
the body or abnormalities in the signal transduction pathway mediated by
NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
trial monitoring and/or the treatment of physiological (heartbeats rate)
or behavioural disorders. NGPCR is useful for identifying compounds
useful in the therapeutic treatment of obesity, inflammation, immune
disorders, diabetes, heart and coronary disease, metabolic disorders and
cancer. Labelled NGPCR nucleotide probes can be used to screen a human
genomic library which is helpful for identifying polymorphisms,
determining the genomic structure of a given locus/allele and designing
diagnostic tests.
Sequence 2526 BP; 469 A; 770 C; 697 G; 587 T; 3 other:

Query Match	Similarity	Score	2524.8	DB	22	Length	2526
Best Local	Similarity	100.0%	Pred.	NO.	0		
Matches	2526	Conservative	0	Mismatches	0	Indels	0
QY	1	atgcgtcctctgcacagcgtctgcctgtgtctgcgcgcgcacagctctcatcttctcgtctgcgtgcgc	60				
Db	1	atgcgtcctctgcacagcgtctgcctgtgtctgcgcgcgcacagctctcatcttctcgtctgcgtgcgc	60				
QY	61	tttgctctgcacatagacaggaagtctctctctgtgctcttcacacctctccgcggagattactctcgt	120				
Db	61	tttgctctgcacatagacaggaagtctctctctgtgctcttcacacctctccgcggagattactctcgt	120				
QY	121	gcagagcgtgtctccctctccatctctgtgctgtctgcacaggtgtagagcacacagacagaggtgac	180				
Db	121	gcagagcgtgtctccctctccatctctgtgctgtctgcacaggtgtagagcacacagacagaggtgac	180				
QY	181	ctgtgtgtgacaggtctctgttagcttcaatgtacgtacagtgctacacctcttcagaggtatgcg	240				
Db	181	ctgtgtgtgacaggtctctgttagcttcaatgtacgtacagtgctacacctcttcagaggtatgcg	240				
QY	241	cttgggtgtgtgagagataaacacaccacccaaagcgcctgcgtctgccacaactacacctgtggtac	300				
Db	241	cttgggtgtgtgagagataaacacaccacccaaagcgcctgcgtctgccacaactacacctgtggtac	300				
QY	301	cagctgtatgtagtgtgtgtgtgtgactctgcacaaatgtgtatgtacacagctgtagtgcctcc	360				
Db	301	cagctgtatgtagtgtgtgtgtgtgactctgcacaaatgtgtatgtacacagctgtagtgcctcc	360				
QY	361	ctgcgcagagcacaacacacacatagagctccaaaggaagacctcttcacattctccctaaaggtg	420				
Db	361	ctgcgcagagcacaacacacacatagagctccaaaggaagacctcttcacattctccctaaaggtg	420				
QY	421	ctggcagtgatgtggcctgcagacacacacacgctgtcgtccacacacagccgctcgtgtgac	480				

Db	421	ctvgcagagatgagatggtccctgcagacacccaacccgctgcgcacacacagccgcgcctgctgtagc	480
Qy	481	cccttcctcgtgtgcaccaagatctaaactatgctgcgcgcacagcagagacgctcacgctgtagaagag	540
Db	481	cccttcctcgtgtgcaccaagatctaaactatgctgcgcgcacagcagagacgctcacgctgtagaagag	540
Qy	541	cagtatccctcttcctgcgcgcacacatccccaatgcacagatcacagatgcagatgcagatgc	600
Db	541	cagtatccctcttcctgcgcgcacacatccccaatgcacagatcacagatgcagatgcagatgc	600
Qy	601	ctgcctgcgcgcagaaaggttcacgggtgtgagcttgagatctctctgtgttgtagacagctgcacacat	660
Db	601	ctgcctgcgcgcagaaaggttcacgggtgtgagatctctctctgtgttgtagacagctgcagatc	660
Qy	661	gggcagcctcagagaggtgcagagcagctgcagagaaacacagccacacgcctgcagggagatctgcattgct	720
Db	661	gggcagcctcagagaggtgcagagcagctgcagagaaacacagccacacgcctgcagggagatctgcattgct	720
Qy	721	ctcaagagacatcatgcgcctctctctcgcacaggtgtagatgcagagagatgcagatgcctcatg	780
Db	721	ctcaagagacatcatgcgcctctctctcgcacaggtgtagatgcagagagatgcagatgcctcatg	780
Qy	781	cgcacacctgcgcgcacagggcccgaggccacccgctcgtgtgtgttttccacagccggagatgtgcgc	840
Db	781	cgcacacctgcgcgcacagggcccgaggccacccgctcgtgtgtgttttccacagccggagatgtgcgc	840
Qy	841	aagggtgttttccagagctccgtgcgtgcgcgcacaaacctgcagcagagagtgtaggtgcgcgcaca	900
Db	841	aagggtgttttccagagctccgtgcgtgcgcgcacaaacctgcagcagagagtgtaggtgcgcgcaca	900
Qy	901	gaagcctgcgcgcctctccacagacatcatcagtgtaggtgcgcgcgcacacagcgcattgcagtag	960
Db	901	gaagcctgcgcgcctctccacagacatcatcagtgtaggtgcgcgcgcacacagcgcattgcagtag	960
Qy	961	gtgcctgcgcgcctgcgcacatccagaaagagaggtgcctccctgcgcgcctcaagagcgcttgtagaagaaagc	1020
Db	961	gtgcctgcgcgcctgcgcacatccagaaagagaggtgcctccctgcgcgcctcaagagcgcttgtagaagaaagc	1020
Qy	1021	tatcccccgggcagagacaaagagggggcccttcagagcctctgcacaaagagggctcctgtgtgcagcagcc	1080
Db	1021	tatcccccgggcagagacaaagagggggcccttcagagcctctgcacaaagagggctcctgtgtgcagcagcc	1080
Qy	1081	aatcacgctctgcagagaaatgcacagctcttcacatgcacacagaaatgcacaaagctcacaaagcc	1140
Db	1081	aatcacgctctgcagagaaatgcacagctcttcacatgcacacagaaatgcacaaagctcacaaagcc	1140
Qy	1141	ttctccatgagttctgtgcctacaacagatccgggctgtgtatgcggtgtgcgcgcacatgcgcctc	1200
Db	1141	ttctccatgagttctgtgcctacaacagatccgggctgtgtatgcggtgtgcgcgcacatgcgcctc	1200
Qy	1201	cacacagcctcctgcgcctgtgcctcctgcagactgtgttcacaggggcgcagctcacacccctgcgag	1260
Db	1201	cacacagcctcctgcgcctgtgcctcctgcagactgtgtgttcacaggggcgcagctcacacccctgcgag	1260
Qy	1261	cttttgtagcagatccacaagagtgatcttcctctctacacacagagacatgtgcgtttaat	1320
Db	1261	cttttgtagcagatccacaagagtgatcttcctctctacacacagagacatgtgcgtttaat	1320
Qy	1321	gacacacagagatcccccctcaatgcataataataattgccttcgggactgcagatgcagaccaaag	1380
Db	1321	gacacacagagatcccccctcaatgcataataataattgccttcgggactgcagatgcagaccaaag	1380
Qy	1381	tgagcccttcacaggtcctccggttccctccacatggtcccccagctccagctacaacataaagag	1440
Db	1381	tgagcccttcacaggtcctccggttccctccacatggtcccccagctccagctacaacataaagag	1440
Qy	1441	accacaaatccaggtgcgcacaggaagagacaaacacaggtgcctcaagctccgtgtgttccacagcgac	1500
Db	1441	accacaaatccaggtgcgcacaggaagagacaaacacaggtgcctcaagctccgtgtgttccacagcgac	1500
Qy	1501	tgctctgcagagggcacaagcgagtggttaacgggttcccatcactgcgtcttgagtggtgtg	1560

Db 1501 tgcctgaaggagccagcgagtggttacgggttcacatcactgctgcttgatgtgtg 1560
QY 1561 ccccttgaggagcttggaacttcctcaacaagaagtacccctcaaatgacacgcttgagg 1620
Db 1561 ccccttgaggagcttggaacttcctcaacaagaagtacccctcaaatgacacgcttgagg 1620
QY 1621 aaagaagagtggaacacttgaaaggaagcagacgcttgctccgacgacgcttgcttg 1680
Db 1621 aaagaagagtggaacacttgaaaggaagcagacgcttgctccgacgacgcttgcttg 1680
QY 1681 gcttgcttgagacacactcttggtgctgctgtaagactaaacgctgctgctgctg 1740
Db 1681 gcttgcttgagacacactcttggtgctgctgtaagactaaacgctgctgctgctg 1740
QY 1741 ctgcttgaggaactgctgctgcttgctgctggaacacccctggtgtagagtgaca 1800
Db 1741 ctgcttgaggaactgctgctgcttgctgctggaacacccctggtgtagagtgaca 1800
QY 1801 gggggccgctgctgcttctctatgctgggtccctgagcagagtgagtgagcctctat 1860
Db 1801 gggggccgctgcttctctatgctgggtccctgagcagagtgagtgagcctctat 1860
QY 1861 ggcctcttgagggaacccacaaagcgctgctgtaagcagccctcttgccct 1920
Db 1861 ggcctcttgagggaacccacaaagcgctgctgtaagcagccctcttgccct 1920
QY 1921 ggttcacacatctcctgctgctgctgagctgctacatccacatatacatcttc 1980
Db 1921 ggttcacacatctcctgctgctgagcagctgctacatccacatatacatcttc 1980
QY 1981 aagtttccacaaagtgacacacatctcaccacgcttggtgctcaaaacacagctgctgc 2040
Db 1981 aagtttccacaaagtgacacacatctcaccacgcttggtgctcaaaacacagctgctgc 2040
QY 2041 ctgcttgatgatcatcagctcagcgccagctgcttctatctgcttaacttgctggtg 2100
Db 2041 ctgcttgatgatcatcagctcagcgccagctgcttctatctgcttaacttgctggtg 2100
QY 2101 tggagccacactgctgctgaggaataacagcgcttccccccttggtgatcttgagtc 2160
Db 2101 tggagccacactgctgctgaggaataacagcgcttccccccttggtgatcttgagtc 2160
QY 2161 acaagaagacacactccctggtctacatactgagcttccctcaaatgagctcctccatc 2220
Db 2161 acaagaagacacactccctggtctacatactgagcttccctcaaatgagctcctccatc 2220
QY 2221 agtgccttgctgctgcaagctacctggtgaagacttgccagagaactacacagagccaa 2280
Db 2221 agtgccttgctgctgcaagctacctggtgaagacttgccagagaactacacagagccaa 2280
QY 2281 tgggtacacttcaagcctgcttcaactctgctcctgagctgcttcttaccacagcgc 2340
Db 2281 tgggtacacttcaagcctgcttcaactctgctcctgagctgcttcttaccacagcgc 2340
QY 2341 agcgttcaagcagcaadtaactgctgctgagcgaacatgagtgagctgagcagcctg 2400
Db 2341 agcgttcaagcagcaadtaactgctgctgagcgaacatgagtgagctgagcagcctg 2400
QY 2401 agcagcgagctcggttggttattcttctgctaagtactgagctgctcctgcccagac 2460
Db 2401 agcagcgagctcggttggttattcttctgctaagtactgagctgctcctgcccagac 2460
QY 2461 ctcaacagacagagcaacttccagagcctcactcaagagctacacagagcgctgcygctcc 2520
Db 2461 ctcaacagacagagcaacttccagagcctcactcaagagctacacagagcgctgcygctcc 2520
QY 2521 acctga 2526
Db 2521 acctga 2526

RESULT 2
AAD19505

ID AAD19505 standard; DNA; 2951 BP.
AC AAD19505;
XX 18-DEC-2001 (first entry)
DT Human novel G-protein coupled receptor DNA with 5' and 3' regions.
XX
XX Human: G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW antisense therapy; signal transduction; behavioural disorder; obesity;
KW heartbeats rate; inflammation; immune disorder; diabetes; cancer;
KW coronary disease; ds.
OS Homo sapiens.
XX WO200172842-A2.
XX 04-OCT-2001.
XX 28-MAR-2001; 2001WO-US09996.
XX 28-MAR-2000; 2000US-192978P.
XX (LEXI-) LEXICON GENETICS INC.
XX Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
PI Zambrowicz B, Sands AT;
XX WPI; 2001-616474/71.
XX
XX Novel isolated polynucleotides encoding human G protein coupled
PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
XX
XX Disclosure; Page 79-80; 80pp; English.
XX
XX The present sequence is human novel G-protein coupled receptor (NGPCR)
CC DNA with 5' and 3' regions. NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences, are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting
CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeats rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.
XX
XX Sequence 2951 BP; 559 A; 888 C; 831 G; 669 T; 4 other;
XX

Query Match 100.0%; Score 2524.8; DB 22; Length 2951;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgctgctctgacgagcctcgctgctgctgctgctgctcattctctgctgctgctg 60
Db 232 atgctgctctgacgagcctcgctgctgctgctgctgctcattctctgctgctgctg 291
QY 61 ttgctgctacatagacgagagcttctctacttaccacccctccgagattactctg 120
Db 292 ttgctgctacatagacgagagcttctctacttaccacccctccgagattactctg 351
QY 121 gcaagcctgttccctctccatctgctgctgctgagtgagacagacccgaggtgacc 180
Db 352 gcaagcctgttccctctccatctgctgctgctgagtgagtgagacccgaggtgacc 411

QY 181 cctgtgacaggtcctgttagcttcaatgagcatggtacacacctcttccaggtatgcg 240
|||||
Db 412 cgtgtgacaggtcctgttagcttcaatgagcatggtacacacctcttccaggtatgcg 471
QY 241 cttgaggttgagagataaacaactccacgcccgtctgcccacaatacaccttggttac 300
|||||
Db 472 cttgaggttgagagataaacaactccacgcccgtctgcccacaatacaccttggttac 531
QY 301 caggtgtatgagtgtgtgtgtgactctgccaatggtatgcccagcttgagagctctcc 360
|||||
Db 532 caggtgtatgagtgtgtgtgtgactctgccaatggtatgcccagcttgagagctctcc 591
QY 361 cgtgcccaggaacaacacacataagagctccaaagagagacctcttccacattcccttaaggtg 420
|||||
Db 592 cgtgcccaggaacaacacacataagagctccaaagagagacctcttccacattcccttaaggtg 651
QY 421 cgtgagagtgatggtggtcctgacagacacaaacgtgtgtgccaacagccgcccgtgtgagc 480
|||||
Db 652 cgtgagagtgatggtggtcctgacagacacaaacgtgtgtgccaacagccgcccgtgtgagc 711
QY 481 ccttctcgtgtgccaatgattagctatgctgccaagagagagagcgtctcagcgtgaaagcgg 540
|||||
Db 712 ccttctcgtgtgccaatgattagctatgctgccaagagagagcgtctcagcgtgaaagcgg 771
QY 541 cagtaacacctcttccctgccaacatacccaatgacaagtaaccaagtgagagacatgtg 600
|||||
Db 772 cagtaacacctcttccctgccaacatacccaatgacaagtaaccaagtgagagacatgtg 831
QY 601 cgtgctgtcagaagatctcggtggtgagacctggaatctctcgtgtgagcagcagtgacagat 660
|||||
Db 832 cgtgctgtcagaagatctcggtggtgagacctggaatctctcgtgtgagcagcagtgacagat 891
QY 661 gggcagctaggggtgagcagcactgagaaacagcagcactggtcagggagctctcatgtct 720
|||||
Db 892 gggcagctaggggtgagcagcactgagaaacagcagcactggtcagggagctctcatgtct 951
QY 721 ttcacagacatcatgacctctctctgcccaggtggtgcgatatgagagatagtcacgtgacctatg 780
|||||
Db 952 ttcacagacatcatgacctctctctgcccaggtggtgcgatatgagagatagtcacgtgacctatg 1011
QY 781 cgcacacttgcccagagccgggcccagccgctcgtgtgtgtttttccagccggcagcttgcc 840
|||||
Db 1012 cgcacacttgcccagagccgggcccagccgctcgtgtgtgtttttccagccggcagcttgcc 1071
QY 841 aggtgtttttccagagtcgtgtgtgtgcagcaaacctgacatgagtgatggtgcgctca 900
|||||
Db 1072 aggtgtttttccagagtcgtgtgtgtgcagcaaacctgacatgagtgatggtgcgctca 1131
QY 901 gaagcttggtccctctccagagcacatcaactgggtgtcccggaatccagcgcatltggatg 960
|||||
Db 1132 gaagcttggtccctctccagagcacatcaactgggtgtcccggaatccagcgcatltggatg 1191
QY 961 ggtgtgtggtgtggtccatccagaagaggtgtgtccctgcgtctgaagcggttttgaagaagcc 1020
|||||
Db 1192 ggtgtgtggtgtggtccatccagaagaggtgtgtccctgcgtctgaagcggttttgaagaagcc 1251
QY 1021 tatgcccgggcaacaagaagagcccttaagccttgccaacaagggcctctgtgtcagaagc 1080
|||||
Db 1252 tatgcccgggcaacaagaagagcccttaagccttgccaacaagggcctctgtgtcagaagc 1311
QY 1081 aatcagctctgcaagatgagcaagcttcaatgacacacagatgcccagaagctcaaaagcc 1140
|||||
Db 1312 aatcagctctgcaagatgagcaagcttcaatgacacacagatgcccagaagctcaaaagcc 1371
QY 1141 tctctcagaagctctgctcaacaagcataccgggctgtgtatgacgttggtccatgtgctc 1200
|||||
Db 1372 tctctcagaagctctgctcaacaagcataccgggctgtgtatgacgttggtccatgtgctc 1431
QY 1201 caacagctctggtgtgtcctcttgagcttcttccaggggcccgaattcaacccctgagcag 1260
|||||
Db 1432 caacagctctggtgtgtcctcttgagcttcttccaggggcccgaattcaacccctgagcag 1491

QY 1261 cttttgagcaagatccacaagggtgcattcctctctcaacaagaagacgttggtcgtttaat 1320
|||||
Db 1492 cttttgagcaagatccacaagggtgcattcctctctcaacaagaagacgttggtcgtttaat 1551
QY 1321 gacacagaagatccctccagtagctataacataattgctgtgagctcgtgatagtgaccag 1380
|||||
Db 1552 gacacagaagatccctccagtagctataacataattgctgtgagctcgtgatagtgaccag 1611
QY 1381 tggaccttaagcgttccctggtccctccacatggtctccaggttccaggttaacataatgag 1440
|||||
Db 1612 tggaccttaagcgttccctggtccctccacatggtctccaggttccaggttaacataatgag 1671
QY 1441 accaanaatccagtgagcaggaagagaacaacaggtgtcctaagtctgtgttccagagac 1500
|||||
Db 1672 accaanaatccagtgagcaggaagagaacaacaggtgtcctaagtctgtgttccagagac 1731
QY 1501 tgtcttgaagggcagcagcagagtggtttaaaggtttccatacactgtcgtcttgaagtgtg 1560
|||||
Db 1732 tgtcttgaagggcagcagcagagtggtttaaaggtttccatacactgtcgtcttgaagtgtg 1791
QY 1561 cctgtgtgggctgagaccttccctccaaagagtgacacctcaacatggtccagcctgtgag 1620
|||||
Db 1792 cctgtgtgggctgagaccttccctccaaagagtgacacctcaacatggtccagcctgtgag 1851
QY 1621 aaagaagatggtgacacctgagggagagccagacctgtctcccgagcagtgatgttttg 1680
|||||
Db 1852 aaagaagatggtgacacctgagggagagccagacctgtctcccgagcagtgatgttttg 1911
QY 1681 gcttgcgtgagagacacacctctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1740
|||||
Db 1912 gcttgcgtgagagacacacctctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1971
QY 1741 cgtctgtgagactgtgacctgt 1800
|||||
Db 1972 cgtctgtgagactgtgacctgt 2031
QY 1801 ggggtccgacctgtgtcttcttaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1860
|||||
Db 2032 ggggtccgacctgtgtcttcttaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2091
QY 1861 ggtctcttgggggaaccccaagacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1920
|||||
Db 2092 ggtctcttgggggaaccccaagacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2151
QY 1921 ggttcaacatctctctgtctgtcctgacagatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1980
|||||
Db 2152 ggttcaacatctctctgtctgtcctgacagatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2211
QY 1981 aagtttccaacaaagttacataatctcaacagcctgtgtgtgtgtgtgtgtgtgtgtgtgt 2040
|||||
Db 2212 aagtttccaacaaagttacataatctcaacagcctgtgtgtgtgtgtgtgtgtgtgtgtgt 2271
QY 2041 cgtttgtatgatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2100
|||||
Db 2272 cgtttgtatgatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2331
QY 2101 tggaccacactgtcgtgtgaggaataccacagcgttcccccacatgtgtgtgtgtgtgtgtgt 2160
|||||
Db 2332 tggaccacactgtcgtgtgaggaataccacagcgttcccccacatgtgtgtgtgtgtgtgtgt 2391
QY 2161 acagagaacaaactccctgtggtctcatactgtgaccttccctcaaatgagcctctctccatc 2220
|||||
Db 2392 acagagaacaaactccctgtggtctcatactgtgaccttccctcaaatgagcctctctccatc 2451
QY 2221 agtgccttggcctgacagctacacctgggttaagagcttgcagagagaactacaagaaggccaa 2280
|||||
Db 2452 agtgccttggcctgacagctacacctgggttaagagcttgcagagagaactacaagaaggccaa 2511
QY 2281 tgytcaacctgaagctgtccttcaactcgtgtctctgtgatacgccttctccacaagccg 2340
|||||
Db 2512 tgytcaacctgaagctgtccttcaactcgtgtctctgtgatacgccttctccacaagccg 2571
QY 2341 agcgttacaagcaggaatacctgtcgtgcgccaacatgatagtgtgtgtgagcagcgtg 2400

```

|||||
Db 2572 agcgctacagcagcgaactacgtcgtcgcccaacatgctggtcgctgagcagctg 2631
Oy 2401 agcagcgcttcgtggtgattttctgcttaagtctacgtgatcctctgcgcgcacac 2460
Db 2632 agcagcgcttcgtggtgattttctgcttaagtctacgtgatcctctgcgcgcacac 2691
Oy 2461 ctcaacacacagcagcactcccaagcctcattcagagctaacagagcgctgcgcctcc 2520
Db 2692 ctcaacacacagcagcactcccaagcctcattcagagctaacagagcgctgcgcctcc 2751
Oy 2521 acctga 2526
Db 2752 acctga 2757

RESULT 3
AADI7516
ID AADI7516 standard; cDNA; 2526 BP.
AC AADI7516;
XX 10-DEC-2001 (first entry)
DE Human taste receptor, hT1R1 cDNA coding sequence.
XX Human: taste-cell-specific G protein-coupled receptor; hT1R1; drug;
KM genetic modulation; pharmaceutical; taste sensation; food industry;
KW chemosensory transduction; ss.
XX Homo sapiens.
OS
XX Key location/Qualifiers
FH 1..2526
FT CDS /*tag= a
FT /product= "Human taste receptor, hT1R1 protein"
FT /transl_except= (pos:820..823, aa:Phe)
FT /note= "This codon has an apparent 1 nucleotide insertion
FT which alters the reading frame"
XX WO200166563-A2.
XX 13-SEP-2001.
PD 13-SEP-2001.
XX 07-MAR-2001; 2001WO-US07265.
PE 07-MAR-2000; 2000US-0187546.
PR 07-APR-2000; 2000US-0195536.
PR 06-JUN-2000; 2000US-0209840.
PR 23-JUN-2000; 2000US-0214213.
PR 17-AUG-2000; 2000US-0226448.
PR 03-JAN-2001; 2001US-0259227.
XX (SENO-) SENOMYX INC.
PA
XX Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
XX WPI: 2001-583267/65.
XX P-PSDB; AAEI0372.
DR
XX New mammalian taste-cell-specific G protein-coupled receptor
PT polypeptides for identifying compounds that modulate taste signaling
PT are useful in food, to modulate the sweet taste of foods or drugs -
XX
XX Claim 1; Page 82-83; 11pp: English.
XX
XX The invention relates to mammalian taste-cell-specific G protein-coupled
XX receptors, T1R and their corresponding cDNA molecules. Taste receptors,
XX T1R are useful for screening compounds which are used to activate or
XX modulate chemosensory transduction, such as taste sensation. The
XX identification and isolation of novel taste receptors and taste
XX signalling molecules allow for new methods of chemical and genetic
XX modulation of taste transduction pathways. The taste modulating

```

CC compounds are useful in pharmaceuticals and food industries to improve
 CC the taste of a variety of consumer products, or to block undesirable
 CC tastes, e.g., in certain pharmaceuticals. T1Rs are also useful in
 CC biochemical assay for identifying tastant (T1R) ligands having binding
 CC specificity for T1R involved in taste signalling. The present cDNA
 CC sequence is human taste-cell-specific G protein-coupled receptor, hT1R1
 CC coding sequence.
 XX
 XX Sequence 2526 BP; 470 A; 770 C; 697 G; 589 T; 0 other;

Query Match 99.0%; Score 2500.2; DB 22; Length 2526;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2521; Conservative 3; Mismatches 1; Indels 2; Gaps 2;

```

Oy 1 atgctgctctcagcagctcgtcgtgctgagctcagctctcattctctgctgagcc 60
Db 1 atgctgctctcagcagctcgtcgtgctgagctcagctctcattctctgctgagcc 60
Oy 61 ttgctgctcagcagcagcagcgtctctcctgagcttcacccctccgagattactctcg 120
Db 61 ttgctgctcagcagcagcagcgtctctcctgagcttcacccctccgagattactctcg 120
Oy 121 gcaagcctgctccctccatctctgctgctgagcgtgagcagacagaccgagtgacc 180
Db 121 gcaagcctgctccctccatctctgctgctgagcgtgagcagacagaccgagtgacc 180
Oy 181 ctgtgtgacaagctctgttagcttcaatagcattgctacaccccttccagagctacgg 240
Db 181 ctgtgtgacaagctctgttagcttcaatagcattgctacaccccttccagagctacgg 240
Oy 241 ctggaggttgagagataaacaaccccgccgctggtcccaataacccctgggtac 300
Db 241 ctggaggttgagagataaacaaccccgccgctggtcccaataacccctgggtac 300
Oy 301 cagcgtatgattgtgtgttctgactctgcacatgtgtatgctgcaagctgagagtgctcc 360
Db 301 cagcgtatgattgtgtgttctgactctgcacatgtgtatgctgcaagctgagagtgctcc 360
Oy 361 ctgccaagggcaacacacatagactccaaggaagaccttccactatccctcagtg 420
Db 361 ctgccaagggcaacacacatagactccaaggaagaccttccactatccctcagtg 420
Oy 421 ctggcagtgattggcctgacagcacaacgctgtctgcacacagccgctctgagc 480
Db 421 ctggcagtgattggcctgacagcacaacgctgtctgcacacagccgctctgagc 480
Oy 481 ccttccctgtgcccattagctatgctgagcgaagcagcgtcagcgctgagacgg 540
Db 481 ccttccctgtgcccattagctatgctgagcgaagcagcgtcagcgctgagacgg 540
Oy 541 cagtaacctcttctctctgagcacaatccccaatgacaagtggagacatgtg 600
Db 541 cagtaacctcttctctctgagcacaatccccaatgacaagtggagacatgtg 600
Oy 601 ctgctgctcagaaagttcggtgagcctgagatctctgtgtgagcagcagtaacacat 660
Db 601 ctgctgctcagaaagttcggtgagcctgagatctctgtgtgagcagcagtaacacat 660
Oy 661 gggcagctaggggtgcaagcacttgagaacacagcgcactggtcagggatctgcatg 720
Db 661 gggcagctaggggtgcaagcacttgagaacacagcgcactggtcagggatctgcatg 720
Oy 721 ttcaagagacatagcctctctctgcccaggtggcgatgagagatgcagtgctcatg 780
Db 721 ttcaagagacatagcctctctctgcccaggtggcgatgagagatgcagtgctcatg 780
Oy 781 cgcacccggccagcgagcgagccacgctgagtggtt-ttttccacccgcaattgac 839
Db 781 cgcacccggccagcgagcgagccacgctgagtggtt-ttttccacccgcaattgac 840
Oy 840 caggtgttttccagctcgtgtgtgctgacaacctgacagtgagtggtgctgcctc 899
Db 840 caggtgttttccagctcgtgtgtgctgacaacctgacagtgagtggtgctgcctc 899

```


Dh	841	caggvgcttttcgagfrcgltgctgagcaaacctgactgccaagfvgltgctgc	900
Qy	900	agaaagcctgggcccctctccagagcaatctggtgtgcocgggataccagcgtatggat	959
Dh	901	agaaagcctgggcccctctccagagcaatctctgggtgtgcocgggataccagcgtatggat	960
Qy	960	ggtctctgggctgtgcaatctccagaagaagggtctgcctctggcctctgaagcggttgaagaagc	1019a
Dh	961	ggtctctgggctgtgcaatctccagaagaagggtctgcctctggcctctgaagcggttgaagaagc	1019b
Qy	1020	ctatgcccgggcsagacaagaaggagcccctcggcctctgcccacaagaagggtcctctgttgcagag	1079a
Dh	1020	ctatgcccgggcsagacaagaaggagcccctcggcctctgcccacaagaagggtcctctgttgcagag	1079b
Qy	1080	caatcagactctgcagagaaatgccaagcttctcatgtrcaacacgatatgccaaagctcaaacg	1139a
Dh	1080	caatcagactctgcagagaaatgccaagcttctcatgtrcaacacgatatgccaaagctcaaacg	1139b
Qy	1140	ctctccatgagtctctgcctccacaagccatccgggtgtgtatctggggtggccatgggct	1199a
Dh	1140	ctctccatgagtctctgcctccacaagccatccgggtgtgtatctggggtggccatgggct	1199b
Qy	1200	ccacagactcctctgggctgtgctccctctgtagctgttctccaaagggtccaggtctaacccctgca	1259a
Dh	1200	ccacagactcctctgggctgtgctccctctgtagctgttctccaaagggtccaggtctaacccctgca	1259b
Qy	1260	gctcttggagcagatccacaagaagtgcattctccttctacaagaagacactgtgctttaa	1319a
Dh	1260	gctcttggagcagatccacaagaagtgcattctccttctacaagaagacactgtgctttaa	1319b
Qy	1320	tgacaacaagaagatcccccctcagtagctataacataatgtcgtgagctggaatggagccaa	1379a
Dh	1320	tgacaacaagaagatcccccctcagtagctataacataatgtcgtgagctggaatggagccaa	1379b
Qy	1380	gtgagacctcaacggtccctctggtctccctccacaatggtctccaggtctccacaataataatga	1439a
Dh	1380	gtgagacctcaacggtccctctggtctccctccacaatggtctccaggtctccacaataataatga	1439b
Qy	1440	gaccacaatccagctgagcagaaagaaagacaacacaggtgtccataagctctgtgttccagcga	1499a
Dh	1440	gaccacaatccagctgagcagaaagaaagacaacacaggtgtccataagctctgtgttccagcga	1499b
Qy	1500	ctgtgcttgaaggggacacagcgagtggtctacgggttccatcacctgctcttgagtggt	1559a
Dh	1500	ctgtgcttgaaggggacacagcgagtggtctacgggttccatcacctgctcttgagtggt	1559b
Qy	1560	gcccctgtgggggctctggagaccttctctcaacaagaagtgcctctcaagaatgcccagcctctgg	1619a
Dh	1560	gcccctgtgggggctctggagaccttctctcaacaagaagtgcctctcaagaatgcccagcctctgg	1619b
Qy	1620	gaaagaaagatgtagcacctctgaggaagaaagacacagacctctccocggcactgtgtgtttt	1679a
Dh	1620	gaaagaaagatgtagcacctctgaggaagaaagacacagacctctccocggcactgtgtgtttt	1679b
Qy	1680	ggctctgagtgagcaaacactcttggtgtgctgtcgtgcagcttaacacgctgtcgtcgtc	1739a
Dh	1680	ggctctgagtgagcaaacactcttggtgtgctgtcgtgcagcttaacacgctgtcgtcgtc	1739b
Qy	1740	gctgcttgaggagactgctggtctgttgcctgggacactaagacaccccttggtgtgaagtacgc	1799a
Dh	1740	gctgcttgaggagactgctggtctgttgcctgggacactaagacaccccttggtgtgaagtacgc	1799b
Qy	1800	agaggggcggcctgtgctcttctatgctgtggccctccctgagcaggttaagtggagcctcta	1859a
Dh	1800	agaggggcggcctgtgctcttctatgctgtggccctccctgagcaggttaagtggagcctcta	1859b
Qy	1860	tgagctcttgggaaaccacaagaagcctgtcgtgtgtcgtacagccagggcctcttgccct	1919a
Dh	1860	tgagctcttgggaaaccacaagaagcctgtcgtgtgtcgtacagggcctcttgccct	1919b
Qy	1920	tggttttacaacattctctgtctgtcgtcagatgtgcatactccaactaatcatcatctt	1979a
Dh	1920	tggttttacaacattctctgtctgtcgtcagatgtgcatactccaactaatcatcatctt	1979b

QY	1980	caaaatttcacacaaaggtaactacatctctacacacagcgtctggtgcacaaacacaaagcgtgcgg	2039
Db	1980	caaaatttcacacacaaaggtaactacatctctacacacagcgtctggtgcacaaacacaaagcgtgcgg	2038
QY	2040	ccctgtctgtatgatcacagctcacaagcggccacagctgtctatctgtcttaacttggctgtgtgt	2099
Db	2040	ccgtgtttgtgatgatcacagctcagcggccacagctgtctatctgtcttaacttggctgtgtgt	2099
QY	2100	gttggaaccacatgcctgctctaggggaataacacagcgtctcccccactctgtgtgagtcttgagt	2155
Db	2100	gttggaaccacactcctctgctctaggggaataacacagcgtctcccccactctgtgtgagtcttgagt	2159
QY	2160	cacagagaccaactccctctgggtgtcttaactgtgcctctctcaaatgtgcctctccat	2219
Db	2160	cacagagaccaactccctctgggtgtcttaactgtgcctctctcaaatgtgcctctccat	2219
QY	2220	cagtgccctttgctgctgcagctactcctgtgtaagacttgcagagaaactacaagagccaa	2279
Db	2220	cagtgccctttgctgctgcagctactcctgtgtaagacttgcagagaaactacaagagccaa	2279
QY	2280	atggtgtacctttagcctgtgtcttctaacttcgtgtgtctgtatcgcgctcttcaacacgac	2339
Db	2280	atggtgtacctttagcctgtgtcttctaacttcgtgtgtctgtatcgcgctcttcaacacgac	2339
QY	2340	caggtgtctaacagcagcgaagtaactgcctgcgtgcgcaacatgagtgtggtgcgtgagagct	2399
Db	2340	caggtgtctaacagcagcgaagtaactgcctgcgtgcgcaacatgagtgtggtgcgtgagagct	2399
QY	2400	gagcagcggcttcggtgtggtatcttctgcctaaagtgtacgtgtacgtctctgcgcgcaga	2459
Db	2400	gagcagcggcttcggtgtggtatcttctgcctaaagtgtacgtgtacgtctctgcgcgcaga	2459
QY	2460	ccctaacagacacagagactctcagagcctccatctacagagactacaagagcgtgtggtgc	2519
Db	2460	ccctaacagacacagagactctcagagcctccatctacagagactacaagagcgtgtggtgc	2519
QY	2520	cacctga 2526	
Db	2520	cacctga 2526	
RESULT 4			
AADI9502			
ID	AADI9502	standard; cDNA: 2292 BP.	
XX	AADI9502;		
AC	18-DEC-2001	(first entry)	
DT			
XX			
DE	Human novel G-protein coupled receptor (NGPCR) cDNA #2.		
XX			
KW	Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;		
KW	antisense-therapy; signal transduction; behavioural disorder; obesity;		
KW	heartbeat rate; inflammation; immune disorder; diabetes; cancer;		
KW	coronary disease; ss.		
XX			
OS	Homo sapiens.		
XX			
XX	Key	Location/Qualifiers	
FH	CDS	1..2292	
FT		/tag- a	
FT		/product- "Human NGPCR protein"	
FT	unsure	85..87	
FT		/tag- b	
FT		/note- "Encodes Ser"	
FT	unsure	880..882	
FT		/tag- c	
FT		/note- "Encodes Ala"	
XX			
XX	W0200172842-A2.		
XX			
XX	04-OCT-2001.		

XX 28-MAR-2001; 2001WO-US09996.
PE 28-MAR-2000; 2000US-192978P.
PR 28-MAR-2000; 2000US-192978P.
XX (LEXI-) LEXICON GENETICS INC.
PA
PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
DR WPI: 2001-616474/71.
DR P-PSDB: AAE11970.
XX
XX Novel isolated polynucleotides encoding human G protein coupled
PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
XX
XX Claim 4; Page 74-75; 80pp; English.

XX The present sequence is a cDNA encoding human novel G-protein coupled
CC receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences, are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting
CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeats rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.
XX
XX

Sequence 2292 BP; 437 A; 690 C; 638 G; 524 T; 3 other:

Query Match 90.7%; Score 2290.8; DB 22; Length 2292;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 225 atgcgcttgagggtgagagataaacaactccacgcccgtgctcccaatcacctg 294
DB 1 atgcgcttgagggtgagagataaacaactccacgcccgtgctcccaatcacctg 60
QY 295 gggtacagctgtatgattgtgttctgactctgccaatgtttatgccaagctgagtg 354
DB 61 gggtacagctgtatgattgtgttctgactctgccaatgtttatgccaagctgagtg 120
QY 355 ctctccctgcagaggaac 414
DB 121 ctctccctgcagaggaac 180
QY 415 acggtgctgcagagtgattgggctgcagacagacacacacacacacacacacacacac 474
DB 181 acggtgctgcagagtgattgggctgcagacacacacacacacacacacacacacacac 240
QY 475 ctgaaccccttccttccttccttccttccttccttccttccttccttccttccttccttc 534
DB 241 ctgaaccccttccttccttccttccttccttccttccttccttccttccttccttccttc 300
QY 535 aagcggacagatccctcttccttccttccttccttccttccttccttccttccttccttc 594
DB 301 aagcggacagatccctcttccttccttccttccttccttccttccttccttccttccttc 360
QY 595 atgtgctgcgtgcgcagaaagtgcgtgcagtgatctctctgttgcagcagtgac 654
DB 361 atgtgctgcgtgcgcagaaagtgcgtgcagtgatctctctgttgcagcagtgac 420

QY 655 gactatggcagcttaggggtgacggcacttgagaaacacagccactggtcaggggagctgc 714
DB 421 gactatggcagcttaggggtgacggcacttgagaaacacagccactggtcaggggagctgc 480
QY 715 atgtcttcaagacatcatgcccctctctgcacagtggtgcagatgagatgagtgatgc 774
DB 481 atgtcttcaagacatcatgcccctctctgcacagtggtgcagatgagatgagtgatgc 540
QY 775 ctcatgcccacttgcccagggcgagggcagacacgctcggtgtgtttttccagccggcag 834
DB 541 ctcatgcccacttgcccagggcgagggcagacacgctcggtgtgtttttccagccggcag 600
QY 835 ttggcagggtgtttttcgaatcgctgtgtgcagaccactgactgtgcaaggtgtgggtc 894
DB 601 ttggcagggtgtttttcgaatcgctgtgtgcagaccactgactgtgcaaggtgtgggtc 660
QY 895 gcttcagaagccttgaggcctctccagcacatcatctgggtgcccggatccaagcact 954
DB 661 gcttcagaagccttgaggcctctccagcacatcatctgggtgcccggatccaagcact 720
QY 955 ggatgtgctggtggcgtggtccatccaagaagggctgtccctgacctgaagcgtttgaa 1014
DB 721 ggatgtgctggtggcgtggtccatccaagaagggctgtccctgacctgaagcgtttgaa 780
QY 1015 gaagcctatgcccgggcagagaagggcccttaggctctgcacaaaggcctctgtgac 1074
DB 781 gaagcctatgcccgggcagagaagggcccttaggctctgcacaaaggcctctgtgac 840
QY 1075 agcacaatcagctctgacagaagaatgccaagcttcatgacacacagatgcccagctc 1134
DB 841 agcacaatcagctctgacagaagaatgccaagcttcatgacacacagatgcccagctc 900
QY 1135 aaagcctctccatgaatgtctgcttacaacagcatccgggctgtgtatgaggtgacct 1194
DB 901 aaagcctctccatgaatgtctgcttacaacagcatccgggctgtgtatgaggtgacct 960
QY 1195 ggctccacacagctcctggtgctgtgctctgagctgttccagggcggaatccaccc 1254
DB 961 ggctccacacagctcctggtgctgtgctctgagctgttccagggcggaatccaccc 1020
QY 1255 tggcagcttttgagacagatccacaaaggtgcatcttccttcaacaaagacgtgtgagc 1314
DB 1021 tggcagcttttgagacagatccacaaaggtgcatcttccttcaacaaagacgtgtgagc 1080
QY 1315 tttatgacaagaagatccctcctcagtgactataacataatgtcgtgagcaggaatgga 1374
DB 1081 tttatgacaagaagatccctcctcagtgactataacataatgtcgtgagcaggaatgga 1140
QY 1375 cccaagtgcaccttcacaggtcctcggtctccctcacatggtccagttcagcctaacaata 1434
DB 1141 cccaagtgcaccttcacaggtcctcggtctccctcacatggtccagttcagcctaacaata 1200
QY 1435 aatgagccaataatccagtgagcaggaaggaacacacaggtgcttaagtctgtgttcc 1494
DB 1201 aatgagccaataatccagtgagcaggaaggaacacacaggtgcttaagtctgtgttcc 1260
QY 1495 agcagctgtcttgaagggcagcagcagtggttgaaggtttccatcacctgcgtcttgag 1554
DB 1261 agcagctgtcttgaagggcagcagcagtggttgaaggtttccatcacctgcgtcttgag 1320
QY 1555 tgtgtccctgtggggctgagaccttccctcaacaagagtgacacttaagatgagcagcct 1614
DB 1321 tgtgtccctgtggggctgagaccttccctcaacaagagtgacacttaagatgagcagcct 1380
QY 1615 tgtggaagaagagtgagacacttgaggaagcagacactgtctcccgacgtgtggtg 1674
DB 1381 tgtggaagaagagtgagacacttgaggaagcagacactgtctcccgacgtgtggtg 1440
QY 1675 ttttggcttggcgtgagcacaaccttgggtgctgcgtgcagcagcacaacagctgtgtc 1734
DB 1441 ttttggcttggcgtgagcacaaccttgggtgctgcgtgcagcagcacaacagctgtgtc 1500
QY 1735 ctgctgcgtctggagcagctgctgttctgtcctgcacactagacacccctgtgtgtgag 1794

QY	550	tcltctctgagcacatccccaattgcaagatcacaggtgtagagacatgvtgctgctg	609
Db	358	tcltctctgagcacatccccaattgcaagatcacaggtgtagagacatgvtgctgctg	417
QY	610	cagaagatctcggtgtagaacctgagatctctctggtttgtagaagatgagatgtagacgta	669
Db	418	cagaagatctcggtgtagaacctgagatctctctggtttgtagaagatgagatgtagacgta	477
QY	670	ggaggtgcaagagcactgtagaagaaacacagagccactgtgtcagaggtatctgcatgtcttcaagagac	729
Db	478	ggaggtgcaagagcactgtagaagaaacacagagccactgtgtcagaggtatctgcatgtcttcaagagac	537
QY	730	atacatgcccctctctctgcagaggtggtgagatgtagagagatgtagatgtccataatgcgcacatg	789
Db	538	atacatgcccctctctctgcagaggtggtgagatgtagagagatgtagatgtccataatgcgcacatg	597
QY	790	gcccagaagccggggccaaacccgctggtgtgtgttttttccagagccggaatttgccaaaggtt	849
Db	598	gcccagaagccggggccaaacccgctggtgtgtgtgttttttccagagccggaatttgccaaaggtt	657
QY	850	ttcagatgcacgttggtgtgtgtacccaacacccggaactgtgtagaaggtgtgtgtgtcgtccacgaagcctg	909
Db	658	ttcagatgcacgttggtgtgtgtacccaacacccggaactgtgtagaaggtgtgtgtgtcgtccacgaagcctg	717
QY	910	gcccctctccagagcacaatcatgctggtgtgcccgggattccagcgcattgtggtgtgctggtgc	969
Db	718	gcccctctccagagcacaatcatgctggtgtgcccgggattccagcgcattgtggtgtgctggtgc	777
QY	970	gttgagcattccagaagaagaggtgtgtccctctggtccgtgaagagcgttttgaaagaagcctatggtccgg	1029
Db	778	gttgagcattccagaagaagaggtgtgtccctctggtccgtgaagagcgttttgaaagaagcctatggtccgg	837
QY	1030	gcacaacaaagagagcccccctaaagcctcttgccacaaagagccctctgtgtgacagacgaatcaagctc	1089
Db	838	gcagaacaaagagagcccccctaaagcctcttgccacaaagagccctctgtgtgacagacgaatcaagctc	896
QY	1090	tgccagagaataatgccaagatcttccatgtrcacacacagatgcccagaagcctcaaaagcctctctcatg	1149
Db	897	tgccagagaataatgccaagatcttccatgtrcacacacagatgcccagaagcctcaaaagcctctctcatg	956
QY	1150	agttctctctcaacaacagatacccggtgtgtgtatgctggtgcccataagagcctccacaagctc	1209
Db	957	agttctctctcaacaacagatacccggtgtgtgtatgctggtgcccataagagcctccacaagctc	1016
QY	1210	cttgagctgtgctctctctgtagagctgtgttccaaaggggcccagatctcaacccccttggaagcatttgtag	1269
Db	1017	cttgagctgtgctctctctgtagagctgtgttccaaaggggcccagatctcaacccccttggaagcatttgtag	1076
QY	1270	cagatccacaagaatgcatcttctctctctacacaagaagacacatgtgtgcgttttaatgacaacaga	1329
Db	1077	cagatccacaagaatgcatcttctctctctctacacaagaagacacatgtgtgcgttttaatgacaacaga	1136
QY	1330	gattccctctgaatgcatataataaataattgtgcctggtgagatctgtagaagcccaagtgtgaccttc	1389
Db	1137	gattccctctgaatgcatataataaataattgtgcctggtgagatctgtagaagcccaagtgtgaccttc	1196
QY	1390	acagttctctggttccctccaatgtgtccagttccagctcgaataaataatgtagaccaaaatc	1449
Db	1197	acagttctctggttccctccaatgtgtccagttccagctcgaataaataatgtagaccaaaatc	1256
QY	1450	cagtgtgcacaggaagaaggaacaacacaggtgtgacctaaagtctgtgtgttccagagctgtgtcttga	1509
Db	1257	cagtgtgcacaggaagaaggaacaacacaggtgtgacctaaagtctgtgtgttccagagctgtgtcttga	1316
QY	1510	gggtgcacagcagatgtgttaaggggttccatcaatctgctgcttggaggtgtgtgcccctgtggg	1569
Db	1317	gggtgcacagcagatgtgttaaggggttccatcaatctgctgcttggaggtgtgtgcccctgtggg	1376
QY	1570	gctgtggagccttctccacaagaaggtgtgaccttcaatagatgccaagtccctgtgtgtggaagaagaag	1629
Db	1377	gctgtggagccttctccacaagaaggtgtgaccttcaatagatgccaagtccctgtgtgtggaagaagaag	1436
QY	1630	tggtgacaccttgaggaagccagagaccttccctccggccagctgtgtgttttttgcttggcttgggt	1689

Db	1437	ttgggaacctgagggaaagccagacctgccttcocggcagcgtgagtgatllttctggcttgcgt	1496
Qy	1690	gagcaacacctctctgggtgctgctcggcgcaagcctaaacagctctctgctctgctctgtcttgg	1749
Db	1437	gagcaacacctctctgggtgctgctcggcgcaagcctaaacagctctctgctctgctctgtcttgg	1556
Qy	1750	actgctgctcgttcttgctcgtcagccactagacaacccctctgtgtgtagtgctcagagggggccgc	1809
Db	1557	actgctgctcgttcttgctcgtcagccactagacaacccctctgtgtgtagtgctcagagggggccgc	1616
Qy	1810	ctgtgtcttcttctatgctcgggctctccctggcagcaggtgagtgagcagcctctatgctctctt	1869
Db	1617	ctgtgtcttcttctatgctcgggctctccctggcagcaggtgagtgagcagcctctatgctctctt	1676
Qy	1870	gggggaaccacaagagccttgctgtctgtctacgcgaagggccctcttttggcccttgattccac	1929
Db	1677	gggggaaccacaagagccttgctgtctgtctacgcgaagggccctcttttggcccttgattccac	1736
Qy	1930	atctctcgtctcgtcgtcgtacagttctgcctcatcttccaaactaatcatcatcttcaagtttcc	1989
Db	1737	atctctcgtctcgtcgtcgtacagttctgcctcatcttccaaactaatcatcatcttcaagtttcc	1766
Qy	1990	accgaaggtacctataattctaacagccttgggtctccaacacacaggtgctgctgttttg	2049
Db	1797	accgaaggtacctataattctaacagccttgggtctccaacacacaggtgctgctgttttg	1856
Qy	2050	atgatacagctcagcggcccgacgctgcttatctgtctaaactctggctctgtgtgtggagcccca	2109
Db	1857	atgatacagctcagcggcccgacgctgcttatctgtctaaactctggctctgtgtgtggagcccca	1916
Qy	2110	ctgcctcgtctagggataaccagcgccttcccccactctgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2169
Db	1917	ctgcctcgtctagggataaccagcgccttcccccactctgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1976
Qy	2170	aactccctgggctcttcaactgagcgtcttcctctataatgagcgtcctctcttcaatgagtgctt	2229
Db	1977	aactccctgggctcttcaactgagcgtcttcctctataatgagcgtcctctcttcaatgagtgctt	2036
Qy	2230	gctctgcaagctcactctgggttaagagacttgcgcagagaaactaacagagcgaatgtygtacc	2289
Db	2037	gctctgcaagctcactctgggttaagagacttgcgcagagaaactaacagagcgaatgtygtacc	2096
Qy	2290	ttaagctctgctcttcaactctgcctgtgctcgtgctcgtctcttcaacagcgccagcgtctac	2349
Db	2097	ttaagctctgctcttcaactctgcctgtgctcgtgctcgtctcttcaacagcgccagcgtctac	2156
Qy	2350	gacggcgaagtactcgtcctgcggccacaatgatatgctctgggctgagcagccttgacggcgcc	2409
Db	2157	gacggcgaagtactcgtcctgcggccacaatgatatgctctgggctgagcagccttgacggcgcc	2216
Qy	2410	ttcgtgtgtgtattttctgacctaaagtgtactgtatctcttgcgcgcgcagagacttcaacgccc	2469
Db	2217	ttcgtgtgtgtattttctgacctaaagtgtactgtatctcttgcgcgcgcagagacttcaacgccc	2276
Qy	2470	acagagcacttcacagcgtcctcatctcagagactaacagagagcgtctggtctccacctga	2526
Db	2277	acagagcacttcacagcgtcctcatctcagagactaacagagagcgtctggtctccacctga	2333
RESULT	6		
AAZ50744			
ID	AAZ50744	standard; cDNA; 2579 BP.	
XX	AAZ50744;		
AC			
XX			
DT	31-MAY-2000	(first entry)	
XX			
DE		Mouse sensory transduction G-protein coupled receptor-B3 cDNA.	
KX		Mouse; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;	
KM		sensory cell; taste receptor cell; screen; taste modulator;	
XX		pharmaceutical; food industry; taste topographic map; tongue; ss.	

XX	Mus sp.	
OS		
XX		
FH	Key	Location/Qualifiers
FT	CDS	11..2539
FT		/tag= a
FT		/product= "GPCR-B3"
XX		
PN	MO200006592-A1.	
XX		
-PD	10-FEB-2000.	
XX		
PE	27-JUL-1999;	99WO-US17099.
XX		
PR	28-JUL-1998;	98US-0094465.
XX		
PA	(REGC) UNIV CALIFORNIA.	
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PI	Zuker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;	
XX		
DR	WPI; 2000-205451/18.	
DR	P-PDB; AAY45022.	
XX		
PT	New isolated sensory transduction G-protein coupled receptor, useful	
PT	for developing products for use in studying and modulating the taste	
XX	transduction pathway -	
PS	Claim 5; Pages 77-79; 83pp; English.	
XX		
CC	The present sequence is a cDNA encoding a taste cell specific G-protein	
CC	coupled receptor, GPCR-B3 which is involved in sensory transduction.	
CC	This sequence was obtained from mouse circumvallate and foliate papillae.	
CC	GPCR-B3 is specifically expressed in foliate and fungiform cells, with	
CC	lower expression in circumvallate taste receptor cells of the tongue.	
CC	The present sequence is used to screen compounds that modulate sensory	
CC	signalling in taste cells, especially taste modulators useful in	
CC	pharmaceutical and food industries to customise taste. The sequence	
CC	can also be used as probe for identifying taste cells and	
CC	subsets of taste receptor cells such as foliate, fungiform and	
CC	circumvallate. Such probes are also useful to generate taste	
CC	topographic maps that elucidate the relationship between the taste	
CC	cells of the tongue and sensory neurons leading to taste centres	
CC	in the brain.	
XX		
SQ	Sequence 2579 BP; 492 A; 771 C; 698 G; 618 T; 0 other;	
	Query Match	66.9%; Score 1689.4; DB 21; Length 2579;
	Best Local Similarity	79.6%; Pred. No. 0;
	Matches 1993; Conservative	3; Mismatches 509; Indels 0; Gaps 0
OY	22 cgtgcgagcctcagaacttcattccctctgctgagccttgacctgcatgaacaggag 81	
Ddb		
	35 cgcgtcagcgtcgagctggcgttgcctctctctctgcttgcgtttcaagccaaggaaaga 94	
OY	82 tcttctcgtactaacctcccgcggagattaccttcggcaggcctgttccctccat 141	
Ddb		
	95 tectctccaggtttcagctccctctctggagcttccctctctgcaagcctgttccctccat 154	
OY	142 ttctgctgtctgcaggttgaggcacaagaccggaggtgaaacctgtgtgacaagctctgtagc 201	
Ddb		
	155 gctgactgtcttcaggttgagacacagacctcttgytttaacaagtgttacagactgtgacagc 214	
OY	202 ttcatatgagcatgagcttaccaccttcccaagcttatgagcgttgggttagagataaac 261	
Ddb		
	215 ttcaacgagcatalgcatltaaccttcccaagcatacggtttcaacggttaaggagataaac 274	
OY	262 aactcacagcctctgctgcacaacatcacctctggggttacagctgatgtatgtgtt 321	
Ddb		
	275 aactcacagcctctgcttcccaaatcacacctctggggttatgaactgtatgtatgtgtctca 334	
OY	322 gactctgcacatgtgtatatgacagcgttgagagtgtcttccctctgcaggagaaacacacata 381	

Dd	335	gagttcttccaaatgctcctaaagcaccctgaaaggtgccccgcctccacgaagaagggacagccacctaa	394
Qy	382	gaagctccaaagaagacactcttccacatcttccccctacagctctgcgcgaatgtaatttggctctac	441
Dd	395	gagatgacgaagaaagatctcttcgcaacacactctcccaaaaggttgtagacatactattggcgctgcat	454
Qy	442	agacaccaacgctgtctgcacacacacagccgcgcctctgtagcccttcccttctgtgtcccatgatt	501
Dd	455	aacactgacccacgcgtctgtccacacactgtgcgcctctctgagcccttctctgtatgtccctgtgtc	514
Qy	502	agctatgtaggcacagccagcgtagagcgtctagcggtgaagccgagatattccctcttctctgagc	561
Dd	515	agctatgtaggcgaacagccagctgtaattccatgaigtggaagcgaagatctccgcctctcttctgagc	574
Qy	562	accatcccccaatgtgacaaggtacacaggtgtagaagacacactgtgtctctgtctgtacgaagattcg	621
Dd	575	accatcccccaagctgaatgaatgtaacaggtgtagaaggtatcatatgtgctgctgtgcgaagagctctgc	634
Qy	622	tggagactggaatctctctgtgttgtagcagcagtgcacgactatgtggcagctaggggtgtagcaga	681
Dd	635	tgggtcttgatctcgtctcgtttgtgagacgtatgtgtacatgtagcagcgtctgtgctgtaacaggtc	694
Qy	682	ctggagagaacccagggccactatgttcaagggatctgtgattgtctttaaaggaaataatgaccttc	741
Dd	695	ctggagagagactgtgcacactccacaggtggatctgtgctgtcccttcaaggaacgttggtgtccttc	754
Qy	742	ctctgcacaggtgtgggcgatactgaagagatgacgtgacgtctcatatgcacacactgtgcacccagcg	801
Dd	755	tcgcgcacaggggtgtgaccccaagatgtagcagcgtatgatagtgctgctgtctgtctgcctcagcag	814
Qy	802	gcccacgctcgtgtgtctgttttcttccacgcgagctgtgcagaggtgtgttttccagctcgtg	861
Dd	815	accacacgtgtgctgtgtctctctctctctaaacgcgcactgtgcgtgtgtcttcttccagctctgt	874
Qy	862	gtgtctgtaccaaactgtactgtgcagaaggtgtgtgtgtctgtctcaagagcctgtgggcctctccag	921
Dd	875	gtctcgtgcacaaactgtactgtgcagaaggtgtgtgtgtctgtctccgaagacgtgggcacactccacg	934
Qy	922	caacatactgtgggtgtgcccgcgagatccacagcgcgatttgggaatgtgtctgtggcgtgtgacatacag	981
Dd	935	tacatacacaactgtgcccgcgagatccacaggtgatttgggaacgtgtgtgggtgtgtgacatacag	994
Qy	962	aagagggctgtctccctgtgcctgtcgtgaagcgttctgaagaagcctatgtgcgcggcagacagaagag	1041
Dd	995	cagagaacaagctccctgtgcctgtgaagagattctgaagaatctctatgtctccagagctgtatgtgt	1054
Qy	1042	ggccctagcctcttgcacaaagggtcctctgtgtgtagcagaacatacagctctgtcgaagaaatgc	1101
Dd	1055	gtctcccaagaactgtgcacaaagggttccctgtgtgtagcactaaacagcgtctgcagggagatgt	1114
Qy	1102	caagcttttcagtgacacacagatgtgcacaaagcttccaaagccttcttccatgtatgtctctctac	1161
Dd	1115	caagcttttcaagacatgtagaaatgtcccgcgagcttggagcctcttccatgtgagcgtctctac	1174
Qy	1162	aacgcataccgggtctgtgtatgcgtgtgcccatagtcctccacagcctctgtggtctgtgc	1221
Dd	1175	aatgtgtatgaggtctgtgtatgtctgtgtgtgcccacgcgtctccacagcctctgtgtagtctac	1234
Qy	1222	ctctgtagctgttcttcacaggggcgagttacacctctgtgcacgttcttgagagagatccacaag	1281
Dd	1235	ctctgtagacctgtgtccaaaggccacagttacacccctgtgcagcttcttcaagagagatctaaag	1294
Qy	1282	gtgcaatttccctctacacaggaacagctgtgtgtgtttaaagacacagaaatccctcgt	1341
Dd	1295	gtgaatttctcttccacataaagaagactgtgtgacattctgaagcaagggggagaccccttagt	1354
Qy	1342	agctataacataatgtctgtgagctgtgaatgtgaccccaagtgtgaccttccaggtctcgt	1401
Dd	1355	tattatgacatacatcgcctgtggactgtgaatgtgagccctgaaatgtgaccttcttgaggttaattgt	1414
Qy	1402	tccttccaaatgtctccagttcagcttaacataaatgtagacaaatcccaatccagttgtgcagga	1461

Db	1415	ctgctccatctgctccagfctactcagacataaataagacaacaaatccagctgagcagg	1474
Oy	1462	aagagacaacagctgctctaaagctctgtgtgtctccagcgacgtctctgaaagggcacaacgcga	1521
Db	1475	aagacaatcaggtgctctgtctacgtgtgctacagggactgtctctcgaaaggcacaacaga	1534
Oy	1522	gtgtgtacgggtttccatcaactctgtcttgggtgtgtgtgcccctgtgagggtcgggaacctc	1581
Db	1535	ctgtctacgtgtcttcccaaccctctgtctctcagatgtgacatccctgtgaaagctgtggacattc	1594
Oy	1582	ctcaacaagaagctgacacctctcaatctcagatctccagacctgtggagaaagaaatgtggacattag	1641
Db	1595	ctcaacaacaggtgagctctcaacacctctgcagaccttgtgtgaacaagaaatggtccctctag	1654
Oy	1642	ggaagccagacacctgtctcccgcgacactgtgtgtgttcttggctcttgcgtgtgagacacctct	1701
Db	1655	ggagagctccagcctgtctctctacgcacccgtggaggtctcttgggggtgctacgaacctctct	1714
Oy	1702	tgggtgtgtctgtgcagctaaacacagctctgtctctgtctgtctgtctgtggactgtggcctgt	1761
Db	1715	ctgtgtctactatgaacagctaaacacgtactgtctgtctgtctgtatgtggagactgtgtgcctgt	1774
Oy	1762	cttgcgcggacacacacaccccccttgggtgtgaaggtcaagcgagggtggcgccctgtctctctt	1821
Db	1775	cttgcgcggcgctcttccacaacgcctgtgtgtgaaggtcagcctgggggtgaagctgtgtctctc	1834
Oy	1822	atgtctgtgctctccctctgtgcagacaggttaaigtgcagaccttaatgtctcttcttggggaaccaca	1881
Db	1835	atgtctgtgctctcctgtgtgaaggttggaggttgacgctctcaagctctctctctgggaagccacg	1894
Oy	1882	aggcctgtgtgtctgtctaaagccagagccctcttgcacctgtgttccacacatctctgtctc	1941
Db	1895	gtgcgcgcgctgtgtctgtgcgtgcagccctctctctctcggtgttgcacattctctctcc	1954
Oy	1942	tgtcctgcagttctgcctcatctcaacacaaatcatcttaagaatttccacaagaagttact	2001
Db	1955	tgtctgcacaacccgctctctctccacacgtgtacatcatcttaagaatttctccacaagaagttacc	2014
Oy	2002	acattctcaacacgctctgtgtgtccacaacacagcgtgtgcgcctgttgtgtatgatacagctca	2061
Db	2015	acattctcaacacactgtggcccaacaaacaaatggtgtccgggaatatctgcatactgtcaagctcc	2074
Oy	2062	gcggccagcgtctactctgtcttaacttggcgtgtgtgtgtgtgtgagcccactgtgcctgtag	2121
Db	2075	acggttcattgttctctctgtctctacacgtgtgtcgtgtacatgttgagccccaagccacaag	2134
Oy	2122	gaataacacagcgtctcccccactctgt	2181
Db	2135	gagataacacagcgtctcccccactctgt	2194
Oy	2182	cttaactgtgcctctccctcaaatgtgcctccctccatcagtgagccttgtgcctgtgcagctac	2241
Db	2195	ttctcgtgtgtccttgcagacacaacatccctccctccatcagtaacacttgtgtgtgcagctac	2254
Oy	2242	ctgtgtgaaggaactgtgcagagaacatacaacagagacaaatgtgtacacttcaagcctgtctc	2301
Db	2255	ctgtgtgaaggaactgtgcagagaacatacaacagagacaaatgtgtacacttcaagcctgtctc	2314
Oy	2302	cttaactctgtctcctgt	2351
Db	2315	cttcaactctgtactctgt	2374
Oy	2362	ctgtcctgcgcacaacatgtatgt	2421
Db	2375	cttaccgcggtgtcaatgt	2434
Oy	2422	ttctctgcctaagtgtactgtgtactctctgtgcgcgcacagacctcaacagacagagaagacttc	2481
Db	2435	ttctctccctaagtgtactgtgtactctctgtgcgcgcacagacctcaacagacagagaagacttc	2494
Oy	2482	cagagctccatcctaaggaactacagaagagcgtgcggtctcaactgtga	2526
Db	2495	cagagctccatcctaaggaactacagaagagcgtgcggtgtgaactgtga	2539

```

RESULT 7
AA250743
ID AA250743 standard; cDNA; 2771 BP.
XX
XX AA250743;
XX
XX 31-MAY-2000 (first entry)
XX
DE Rat sensory transduction G-protein coupled receptor-B3 cDNA.
XX
XX Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
KW sensory cell; taste receptor cell; screen; taste modulator;
KW pharmaceutical; food industry; taste topographic map; tongue; ss.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT CDS 57..2579
FT FT /tag="a
FT FT /product="GPCR-B3"
XX
XX PN MO200006592-A1.
XX
XX 10-FEB-2000.
XX
XX 27-JUL-1999; 99WO-US17099.
XX
XX 28-JUL-1998; 98US-0094465.
XX
XX (REGC ) UNIV CALIFORNIA.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
XX
XX WPI: 2000-205451/18.
XX DR P-PSDB; AA45021.
XX
XX New isolated sensory transduction G-protein coupled receptor, useful
PT for developing products for use in studying and modulating the taste
PT transduction pathway -
XX
XX Claim 5; Pages 76-77; 83pp; English.
XX
XX The present sequence is a cDNA encoding a taste cell specific G-protein
CC coupled receptor, GPCR-B3 which is involved in sensory transduction.
CC This sequence was isolated from the IZAP rat circumvallate cDNA library.
CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with
CC lower expression in circumvallate taste receptor cells of the tongue.
CC The present sequence is used to screen compounds that modulate sensory
CC signalling in taste cells, especially taste modulators useful in
CC pharmaceutical and food industries to customise taste. The sequence
CC can also be used as probe for identifying taste cells and
CC subsets of taste receptor cells such as foliate, fungiform and
CC circumvallate. Such probes are also useful to generate taste
CC topographic maps that elucidate the relationship between the taste
CC cells of the tongue and sensory neurons leading to taste centres
CC in the brain.
XX
XX Sequence 2771 BP; 564 A; 797 C; 754 G; 656 T; 0 other;

```

Query Match	Score	1674.4	DB 21	Length 2771
Best Local Similarity	78.9%	Pred. No. 0		
Matches 1990	Conservative	3	Mismatches 529	Indels 0
QY	5	tgctctgcacagcgtcgccgtgcgcgcctgcagcttcacattccctgctgctggccttg	64	
DB	58	tgctcttcggagctgctacccctgctcgtcagcctgagctgagcttgctacacgtcggcctta	117	
QY	65	ccctgcacatgcacagagcttcttcctgcactcaacctcccgagatcactcctgcag	124	

[illegible]

QY	1205	agctctcgggctcgtgcctctcggagcttgcttcaggggcccaggtctacccctcggccgttt	1268
-Db	1258	agctctcgggagtgtaactcctcgtgatactcgtctccagagccccaagctctacccctcggccgtttc	1317
QY	1265	tggagcaatccacaagaagtgcatttccctctacaagaagacactcgtgcgtttaatgtaca	1334
-Db	1318	ttcagcagatactacaagggtgaatttctctctacatgagaaatactcgtgcatttgtagca	1377
QY	1325	acagagatccctcaagtaagctataataatctgcctcgtggaactcgtgaatcggaccagtgga	1384
-Db	1378	acggggagacactctagtgtaactacagacatactgcgctcgggagctcgtgaatcgtgacctgaa	1437
QY	1385	ccctcaacgctccctcgtgtctccccaacatcgtcctcaactcgaactcgaacataaattgagaca	1444
-Db	1438	cccttgagaatactatcgtcctcgtcctcactcgtctccagttcactctgacataataaagaaca	1497
QY	1445	aaatccagtcggcacggagaaagacaaccaagctgcctcaagctcgtgtgtcttcagagacgtc	1504
-Db	1498	aaatccagtcggcacggagaaagacaacatcagaatgcctgtgtcaagtggtacacagagatgc	1557
QY	1505	ttgaaaggcacacagccaggttgtaacgggtttccataacgctcgtccttgaggtgtgcacct	1566
-Db	1558	tggcagggcacccaagagtcgggtctgtgggtctccacaacacgtcgtcttgaggtgtgtgcacct	1617
QY	1565	gtcgggacgtggagaccttccctcaacaagaagtgaaacctctacagatgcagaccctctgtggaaag	1624
-Db	1618	ggagaaagctgggagaccttccctcaacaatggaggttcacatctccagagcctctgtggaaag	1677
QY	1625	aagagtcgggacacctcggagggagaaagccagaacacctgcctcccgacactcgtgtgttttgacct	1684
-Db	1678	aagaaatgggacccaaccaagagagagacatactgcctccacaagcagcagtggaagttcttgacct	1737
QY	1685	tgcgtggaacacacctctctgggtgtgtctgtgcagctacaacgctcgtcgtcgtcgtcgtcgc	1744
-Db	1738	ggcagaaacccaactctcttttggtgtgttaataagcagctaaacgcgtatgtcgtcgtcgtcgcgtg	1797
QY	1745	ttggagactcgtgcgtcttggtcctgtgcacctaagacacctctgtgtgaggttcagcagggg	1804
-Db	1798	ttggagacactcgtgcgtcttggtcctgtgcacatttccacaacactgtaagtgaaagttcagatcggg	1857
QY	1805	gcccgcctgtgtcttctctatgactggagcccccctgtgcacagaatgctgcacacccatactgct	1864
-Db	1858	gtagagctgtgtcttctcctaactcgtggttccctcgttggtgcgggaagttgcagctctatact	1917
QY	1865	tcttctgggaaaccaagaagcctcgtcgtctgtctgaagcagagccctcttgcctctggtc	1924
-Db	1918	tcttctgggagcccaacagttgcgccgtcgtctgtctgtcgaagccctcttctcctcgtggt	1977
QY	1925	tcaacatttctcgttctcgtcctgaacagtttgcgtcatttccaacttaataatcattctcaagt	1984
-Db	1978	ttgcacattctcctcctcgtcctgtgcacataccggtcccttccaactgacatcacaatttcaagtc	2037
QY	1985	tttccacaagaagtacataacttataacttacaacgctcgggttccaaaacaacagcgtctgcgcgt	2044
-Db	2038	tttccacaagaagtgcaccaactctcaacgtcagccctgagggcccaaaaacaatggttgaggtctat	2097
QY	2045	ttgtgataatcagctcagccgagcccaagctgtctatactgtctaactcgtgcgtgtgtgtgtga	2104
-Db	2098	tcgtcatctgtaagctccaagctcatcttggtctatactgtcttcacatagctcgttgtaagtgtga	2157
QY	2105	cccacactcgtcgtcgaaggaaatacacaggtctccccaactcgttgtagtcttgaggtgcagag	2164
-Db	2158	cccacagccccaacaaggaaatacacaggtctccccaactcgttgtagtcttcctgaggtgcagag	2217
QY	2165	agacaaactccctcgtgtctataactagcgtccttcccttacaagaatgcgtccctctccatcaag	2224
-Db	2218	aggtcaactcgtgaggtcttcctcgttggtctttaaccacaacaatctctcctctccatcagta	2277
QY	2225	cccttgctcgtcagctacacctcgggtgaaggaactgtccagagaactacaagagagccaatgtg	2284
-Db	2278	cccttgctcgtcagctacacctcgggtgaaggaactgtccagagaactataatgaagccaatgtg	2337

QY 2285 tcacctcagctgctctcaacttcgtctcgtgacgcctctctcaccagcgcaqg 2344
|||||
Db 2338 tcaactcagctgctcctcaacttcgtatctctgacgtctctcaccatggcgca 2397
QY 2345 tctacagcgcaagtactgctcgtcgccacaatatagtgctggctgagcagctga 2404
|||||
Db 2398 ttaccagcgcaactactgctcgtcgctcaatgctgctgcaaggtgaccacacagcg 2457
QY 2405 gcggcttggtggtggtattcttcgtctaagtgtacgtatctctgcgcgcagcctca 2464
|||||
Db 2458 gcggcttcagcggtactctccccaagtgctatgtatctctgcgcgtccagaactca 2517
QY 2465 acagacagagcactcagcgcctcattcagactacagagcgctgctgcctcacct 2524
|||||
Db 2518 acaatacagaaacttcaagcctccatccaggaactacaagagcgctgcggaactacct 2577
QY 2525 ga 2526
Db 2578 ga 2579

RESULT 8
AAD19503
ID AAD19503 standard; cDNA: 1101 BP.
XX
AC AAD19503;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human novel G-protein coupled receptor (NGPCR) cDNA #3.
XX
KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW antisense-therapy; signal transduction; behavioural disorder; obesity;
KW heartbeats rate; inflammation; immune disorder; diabetes; cancer;
KW coronary disease; ss.
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT CDS 1..1101
FT /tag= a
FT /product= "Human NGPCR protein"
FT unsure 520..522
FT /tag= b
FT /note= "Encodes Ala"
FN MO200172842-A2.
PN 04-OCT-2001.
XX
PD 04-OCT-2001.
XX
PE 28-MAR-2001; 2001MO-US09996.
XX
PR 28-MAR-2000; 2000US-192978P.
XX
PA (LEXI-) LEXICON GENETICS INC.
PI Walke DM, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
XX
DR WPI: 2001-616474/71.
XX
P-PSDB; AAE11971.
XX
PT Novel isolated polynucleotides encoding human G protein coupled
PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
XX
XX Claim 5; Page 76-77; 80pp; English.
XX
PS The present sequence is a cDNA encoding human novel G-protein coupled
CC receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences, are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting

CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeats rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.
SQ Sequence 1101 BP; 219 A; 308 C; 323 G; 250 T; 1 other:

Query Match 39.6%; Score 1001.2; DB 22; Length 1101;
Best Local Similarity 99.6%; Pred. No. 3e-243;
Matches 1004; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 595 atgtgtcgtcgtcgtcagagaagtcgggtggacctgacatctctggttggcagcaatgac 654
Db 1 atgtgtcgtcgtcgtcagagaagtcgggtggacctgacatctctggttggcagcaatgac 60
QY 655 gactatggcagctagggtggtgagcagcactggagaaacagcgcaactggtcagagggtatgc 714
Db 61 gactatggcagctagggtggtgagcagcactggagaaacagcgcaactggtcagagggtatgc 120
QY 715 atgtcttcaagaacatagcctctctctgcccaggtggtgagatgagatgagatgac 774
Db 121 atgtcttcaagaacatagcctctctctgcccaggtggtgagatgagatgagatgac 180
QY 775 ctcatgcccacactgcccagggcgccagcgtctgtgtgtttttccagccggcag 834
Db 181 ctcatgcccacactgcccagggcgccagcgtctgtgtgtttttccagccggcag 240
QY 835 ttggcagaggtgtttttcgaagtcgtgtgtgacacactgacatgcaaggtgtgtgtc 894
Db 241 ttggcagaggtgtttttcgaagtcgtgtgtgacacactgacatgcaaggtgtgtgtc 300
QY 895 gcttcagaagccttggcctctctcagcagacatcactgggtgtgtgtgtgtgtgtgtgt 954
Db 301 gcttcagaagccttggcctctctcagcagacatcactgggtgtgtgtgtgtgtgtgtgt 360
QY 955 gggatgt 1014
Db 361 gggatgt 420
QY 1015 gaagcctatgcccggcgagacagaagggcccttgaagccttgcacaaaggcctctgtgtc 1074
Db 421 gaagcctatgcccggcgagacagaagggcccttgaagccttgcacaaaggcctctgtgtc 480
QY 1075 agcagaacatcagctctgacagaaatgcaagcttcatgtrcacacagatgcccagaatc 1134
Db 481 agcagaacatcagctctgacagaaatgcaagcttcatgtrcacacagatgcccagaatc 540
QY 1135 aaagccttctcattagtttctgtctacaaagataccgggtgtgtgtgtgtgtgtgt 1194
Db 541 aaagccttctcattagtttctgtctacaaagataccgggtgtgtgtgtgtgtgtgt 600
QY 1195 ggcctcacaacagctcctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1254
Db 601 ggcctcacaacagctcctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
QY 1255 tggcagcttttggagcagatccacaaggtgcatcttctcttctacacaagacactgtg 1314
Db 661 tggcagcttttggagcagatccacaaggtgcatcttctcttctacacaagacactgtg 720
QY 1315 ttatagcaaaagagatccctcagtgatataataatgtgctggagcagatgga 1374
Db 721 ttatagcaaaagagatccctcagtgatataataatgtgctggagcagatgga 780

Chr	Start (kb)	End (kb)	Gene	Transcript	Exon	Intron	UTR	Feature	Annotation
1	133.75	133.75	CCNA1	CCNA1	1			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	2			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	3			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	4			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	5			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	6			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	7			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	8			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	9			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	10			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	11			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	12			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	13			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	14			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	15			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	16			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	17			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	18			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	19			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	20			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	21			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	22			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	23			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	24			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	25			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	26			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	27			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	28			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	29			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	30			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	31			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	32			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	33			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	34			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	35			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	36			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	37			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	38			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	39			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	40			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	41			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	42			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	43			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	44			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	45			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	46			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	47			133.75	CCNA1
1	133.75	133.75	CCNA1	CCNA1	48			133.	

```

FT          /number= 6
FT          7772..8191
FT          /tag= n
FT          /number= 7
FT          /cons_splince= (5'site:NO, 3'site:NO)
PN          MO20016563-AZ.
PD          13-SEP-2001.
PE          07-MAR-2001; 2001WO-US07265.
PR          07-MAR-2000; 2000US-0187546.
PR          07-APR-2000; 2000US-0195536.
PR          06-JUN-2000; 2000US-0209840.
PR          23-JUN-2000; 2000US-0214213.
PR          17-AUG-2000; 2000US-0226448.
PR          03-JAN-2001; 2001US-0259227.
XX          (SEMO-) SENOMYX INC.
PA          Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
PI          WPI; 2001-582267/65.
DR          P-PSDB; AAEI0372.
XX          New mammalian taste-cell-specific G protein-coupled receptor
PT          polypeptides for identifying compounds that modulate taste signaling
PP          are useful in food, to modulate the sweet taste of foods or drugs -
PS          Claim 1; Page 78-81; 119pp; English.
XX          The invention relates to mammalian taste-cell-specific G protein-coupled
CC          receptors, T1R and their corresponding cDNA molecules. Taste receptors,
CC          T1R are useful for screening compounds which are used to activate or
CC          modulate chemosensory transduction, such as taste sensation. The
CC          identification and isolation of novel taste receptors and taste
CC          signalling molecules allow for new methods of chemical and genetic
CC          modulation of taste transduction pathways. The taste modulating
CC          compounds are useful in pharmaceuticals and food industries to improve
CC          the taste of a variety of consumer products, or to block undesirable
CC          tastes, e.g., in certain pharmaceuticals. T1Rs are also useful in
CC          biochemical assay for identifying tastant (T1R) ligands having binding
CC          specificity for T1R involved in taste signalling. The present sequence is
CC          human taste-cell-specific G protein-coupled receptor, hT1RI full-length
CC          genomic DNA.
SQ          Sequence 8191 BP; 1695 A; 2332 C; 2162 G; 1902 T; 100 other;

Query Match      36.0%; Score 908.6; DB 22; Length 8191;
Best Local Similarity 99.7%; Pred. No. 1.9e-219;
Matches 930; Conservative 1; Mismatches 0; Indels 2; Gaps 2

OY          1594 gacctctacagatgccagccttctgtgggaagaagaattggtgcacttaagggaaccagacc 1653
              |||||
DB          6841 gacctctacagatgccagccttctgtgggaagaagaagtggcaccttgaggaagccagacc 6900
              |||||
OY          1654 tgcctcccgcgacgctgttgttttggttggttcgctgaaacaacctcttggtgctgctc 1713
              |||||
DB          6901 tgcctcccgcgacgctgttgttttggttggttcgctgaaacaacctcttggtgctgctc 6960
              |||||
OY          1714 gaagctaacaagctgtctgtctgtctgtcgtcttggaagcgtgcgctgttctgtctgcac 1773
              |||||
DB          6961 gcagctaacaagctgtctgtctgtctgtcgtcttggaagcgtgcgctgttctgtctgcac 7019
              |||||
OY          1774 ctagaacacctgtgtgtgaagtcacaagggggcgcgctgtgtcttctaatagtggctccc 1833
              |||||
DB          7020 ctagaacacctgtgtgtgaagtcacaagggggcgcgctgtgtcttctaatagtggctccc 7079
              |||||
OY          1834 ctggcgaagcaggtgaatggcagcctctatagtctcttcttggggaaaccaagaagcctgtgc 1893
              |||||
DB          7080 ctggcgaagcaggtgaatggcagcctctatagtctcttcttggggaaaccaagaagcctgtgc 7139

```


OY	1894	ttgttaagcgaagcgcctcttgcccttggtttcaacaattctctctcgtcctgaacat	1953
Db	7140	ttgttaagcgaagcgcctcttgcccttggtttcaacaattctctcgtcctgaacat	7199
OY	1954	cgctcatctccaactaatacatcatcttcaagttctccacaagaagtacattcatcacac	2013
Db	7200	cgctcatctccaactaatacatcatcttcaagttctccacaagaagtacattcatcacac	7259
OY	2014	gctcgtgggtccaacaacacggtctgtgcctcgttttgtatgatctagtcagcgccagctg	2073
Db	7260	gctcgtgggtccaacaacacggtctgtgcctcgttttgtatgatctagtcagcgccagctg	7319
OY	2074	cttactcgtctaaacttggcgtctgtgtgtgtgaaccacatgcctgtcagtaggaatacacagcg	2133
Db	7320	cttactcgtctaaacttggcgtctgtgtgtgtga -cccactgcctgtcagtaggaatacacagcg	7378
OY	2134	ttcccccacatctgtatgatctgtatgacagaagaccacatccctcttggtgttcatactgccc	2193
Db	7379	ttcccccacatctgtatgatctgtatgacagaagaccacatccctcttggtgttcatactgccc	7438
OY	2194	ttcctctacaagatggcctcctcttcatacatgtccttctgctctgtagtactactggttaagagac	2253
Db	7439	ttcctctacaagatggcctcctcttcatacatgtccttctgctctgtagtactactggttaagagac	7498
OY	2254	ttgccaagagaacatcaacaacgagggccaaatctgttcaactcttaagcctgctcttcaactctg	2313
Db	7499	ttgccaagagaacatcaacaacgagggccaaatctgttcaactcttaagcctgctcttcaactctg	7558
OY	2314	ttctctgatcgcctctcttctacacacgcccagcgcgtctacgacgccaagtactcctgtcgcgc	2373
Db	7559	ttctctgatcgcctctcttctacacacgcccagcgcgtctacgacgccaagtactcctgtcgcgc	7618
OY	2374	aaacttgatctggcctcgtcgaagacacgctcgaagcagcgttcgtgtgggtattctctgcctag	2433
Db	7619	aaacttgatctggcctcgtcgaagacacgctcgaagcagcgttcgtgtgggtattctctgcctag	7678
OY	2434	tgctacagtatctcctctgcgcgcccaagacccctcaacagacagacgaatcttcagagccttcaat	2493
Db	7679	tgctacagtatctcctctgcgcgcccaagacccctcaacagacagacgaatcttcagagccttcaat	7738
OY	2494	caggaactacaacagagcgctcgtcgcgtctcaacctga	2526
Db	7739	caggaactacaacagagcgctcgtcgcgtctcaacctga	7771

RESULT	10
AAD19504	
ID	AAD19504 standard; cDNA; 705 BP.
XX	
XX	
AC	AAD19504;
XX	
DT	18-DEC-2001 (first entry)
XX	
DE	Human novel G-protein coupled receptor (NGPCR) cDNA #4.
XX	
KW	Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW	antisense-therapy; signal transduction; behavioural disorder; obesity;
KW	heartbeats rate; inflammation; immune disorder; diabetes; cancer;
KW	coronary disease; ss.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	Location/Qualifiers
CDS	1..705
FT	/*tag= a
FT	/product= "Human NGPCR protein"
XX	
PN	W0200172842-A2.
XX	
PD	
XX	04-OCT-2001.
XX	
PF	28-MAR-2001; 2001MO-US09996.

XX 28-MAR-2000; 2000US-192978P.
PR
XX
XX (TEXT-) LEXICON GENETICS INC.
PA
XX
XX
PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abdul A;
PI Zambrowicz B, Sands AT;
XX
XX WPI: 2001-616474/71.
DR P-PSDB; AAE11972.
DR
XX
XX Novel isolated polynucleotides encoding human G protein coupled
PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
PT expressed NGPCRs for diagnosis of disease, and as probes or primers
XX
XX Disclosure, Page 78; 80pp; English.

CC The present sequence is a cDNA encoding human novel G-protein coupled
CC receptor (NGPCR). NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences, are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting
CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeats rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.

Sequence 705 BP; 130 A; 231 C; 170 G; 173 T; 1 other;

Query Match	27.9%;	Score 704.6;	DB 22;	Length 705;
Best Local Similarity	100.0%;	Pred. No. 2.8e-168;		

QY	1822	atgcggagctccctcgtgacgagaaagatgtagcaacccctaaagctctcttttggggaaaccaca	1881
Db	1	atgcggagctccctcgtgacgagaaagatgtagcaacccctaaagctctctcttggggaaaccaca	60
QY	1882	aggctgtgtgtctgctacgacgagccctcttgcctctgtgtttcaacatctctctgtcc	1941
Db	61	aggctgtgtgtctgctacgacgagccctcttgcctctgtgtttcaacatctctctgtcc	120
QY	1942	tgcctgacagttgcgtcatctccaactaaatcatcatcttcaagttttccaacaaagtacct	2001
Db	121	tgcctgacagttgcgtcatctccaactaaatcatcatcttcaagttttccaacaaagtacct	180
QY	2002	acattctacacagcgtctgtgtccaaaacacacaggtgtctgtgccttctgtgtatgtacgtcta	2061
Db	181	acattctacacagcgtctgtgtccaaaacacacaggtgtctgtgccttctgtgtatgtacgtcta	240
QY	2062	ggcgccacgagctgcgtatctgtctaaacttgcgtgcgtggtgtgtgaaccccaatgctcttag	2121
Db	241	ggcgccacgagctgcgtatctgtctaaacttgcgtgcgtggtgtgtgaaccccaatgctctcttag	300
QY	2122	gaataccagcgtctccccaatctgtgtatgtcttggatgtgacagaaagacaacacctccctgggc	2181
Db	301	gaataccagcgtctccccaatctgtgtatgtcttggatgtgacagaaagacaacacctccctgggc	360
QY	2182	ttcatatctgacctctcctctacaaatgagcctcctctccatcaatgtgcttttgcctgtcacgtac	2241
Db	361	ttcatatctgacctctcctctacaaatgagcctcctctccatcaatgtgcttttgcctgtcacgtac	420
QY	2242	ctgggttaagaccttgcacagagaacatacaacagagcccaaatggttcaacttcaagcctgtctc	2301

RESULT 12
 AAS46390/c ID AAS46390 standard; DNA: 5491 BP.
 XX AC AAS46390;
 DT 18-DEC-2001 (first entry)
 XX Tumour suppressor gene derived chemically modified sequence #112.
 DE Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
 KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KM cytosine methylation; ds.
 CS Homo sapiens.
 XX
 XX WO200168912-A2.
 XX
 XX 20-SEP-2001.
 PD
 PF 15-MAR-2001; 2001WO-EP02955.
 XX
 XX 15-MAR-2000; 2000DE-1013847.
 XX 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIDENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K;
 PI
 DR WPI; 2001-602752/68.
 XX
 PT Fragments of chemically modified genes associated with tumour suppressor
 PT genes and oncogenes, useful in designing primers and probes for
 PT analysing diseases associated with cytosine methylation state e.g.
 PT cancer -
 PS
 XX Claim 1; SEQ ID NO 112; 27pp: English.
 XX
 XX The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphate, of genes associated with tumour suppression and
 CC oncogenes having a sequence taken from 536 (actually 533 since
 CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
 CC (SS) and sequences complementary to (SS). The nucleic acid may be a
 CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
 CC form part of a set of probes for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the
 CC differences serving as basis for diagnosis and/or prognosis events which
 CC are disadvantageous to patients. The present sequence is one of the
 CC 533 genomic sequences derived from tumour suppressor genes and
 CC oncogenes. Sequences with even numbered Seq ID numbers are the
 CC complementary sequence of the corresponding odd numbered sequence (e.g.
 CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
 CC is missing).
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 S0 Sequence 5491 BP; 1262 A; 176 C; 1573 G; 2480 T; 0 other;

	Best Local Similarity	76.7%	Pred. No. 6.8e-138	
	Matches 715:	Conservative 1:	Mismatches 216:	Indels 0: Gaps 0:
QY	1595	acccttaacagatgcagcctctgtggaagaagaagtgggacaccttagaggaagccagacct	1654	
Db	1889	ACCTCTACAAATACCAACCTTTATAAAAAAATAAACAACCTTAAAAAAACCAACCT	1830	
QY	1655	gtctcccgagcactgtgtgttcttggtcttgctgtgagacacacctctgggtgcctcgg	1714	
Db	1829	ACTTCCCGGCACTATTAATATTTTAACTTTAGTAAGTAACACACCTTTAAATATCTACGTA	1770	
QY	1715	cagcttaacacgctgtcctgcctgcctgcctctgtggagcgcctgtgcctgtctgcgcac	1774	
Db	1769	CAACTTAACGCGTACTACTACTACTACTACTTAATAAAGTAACTTAACCTATTACTTAACAC	1710	
QY	1775	tagaacaccctctgtgtgagtgacgaagggcgccctctgtcttcttaatgtcggctccc	1834	
Db	1709	TAAACACCCCTTAAATTAATAATCAACAAAAAAGCGCCTATTCTTATTAATAAATCTGCC	1650	
QY	1835	tggcagcagtgatgtgagcaccctatgtcctcttctggggaaccacaagcctgtgtct	1894	
Db	1649	TAAACAACAATATATAACAACCTTAACTTAACTTTTAAAAAAGCCCAAAACCTAGGTACT	1590	
QY	1895	tgtctacgacagcgccctcttgcctctgtgtttcacatctcctctgcctgcctgcagcttc	1954	
Db	1589	TACTTAAGCAAAACCTCTTTACCTTAATTTTACCATCTCTTATCTTAACCTTAACAAATTC	1530	
QY	1955	gtctatctccaactaatcatcatcttccttaagttcttcacaaaggtacatcatctcaag	2014	
Db	1529	GCATATCCAACTAATCATCATCTTCAAAATTTTCCACCAAAATTAACCTAATCTACACAG	1470	
QY	2015	cctgtgtgccaaaacacagcgtgtcgtgcctgttctgtatgatcagctcagcgccacgtgc	2074	
Db	1469	CCTAAATCCAAAACCAACGATACTAATCCTATTTATATATTAACCAACGACCAACATAC	1410	
QY	2075	ttatctgtcttaactgtgcctgtgtgtgttgagaccacactcctctctagggaaatacagcgt	2134	
Db	1409	TTATCTATTAACCTTAACCTTAATATAATAAACCCTACTACTTAATAAATAATACCAACGT	1350	
QY	2135	tcccccacttgtgtatgtctttagtgacgaagagacaaactcccttggccttcatatctgcct	2194	
Db	1349	TCCCCCATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1290	
QY	2195	tcccttaacaatgtgcctcctctccatcagtgctcttgcctgcagctactcgtggtgaagct	2254	
Db	1289	TCTCTACAAATTAACCTCCTCTCCATCAATACCTTTACTACTAATTAATTAATTAAT	1230	
QY	2255	tgcagagagactacaagagcgccaatagtctactcttaagcctctgtcttcaacttcgtgt	2314	
Db	1229	TACCAAAAAACATACAAACGAACCAATATATATACCTTCAACCTACTCTTCAACTTCGTAT	1170	
QY	2315	ccttgatgcctctcttctcaacacgcgcagcgtcttaagacgcgaagtacatcctctgcgcga	2374	
Db	1169	CCTAAATGCGCTTCTTTCACACAGACCAACAGCTTAGAGAGCAAAATACCTACTAGACACA	1110	
QY	2375	acatgatgcctgtggtcgtgagcagccttgagcagcgcctctcgtgtgtatcttctgcstaagt	2434	
Db	1109	ACATATAAATTAACCTAACTTAACCAACCTTAACCAACAGATTCGATTAATTTTCTACTTAAT	1050	
QY	2435	getaagtatcctctctgcgcgcagacacttcaagaacagagagacttccagcctcatctc	2494	
Db	1049	ACTAGTAATTCCTTACCGCCCAAAACCTCAACAAACAAACACATTCACAAACCTCATTC	990	
QY	2495	agactacacagcgcctgtcgcctcaacttga	2526	
Db	989	AAAATACACGAACGCTACGACTCCACTTA	958	

[illegible][illegible]

D	b	1463	atgctctctgtatcaaccccaacaacacgcttccctctctccatgtgtctcaagagctcgc	1522
Q	y	1505	ttgaaggagccacagcgatgtgttaagtggtttccatcactgcgtctttgaagtgtgcctc	1554
D	b	1523	agccaaagagcaaatgtaaaaagctctgtgtgctccacccctgtgtctctcgaatgtctgaa	1582
Q	y	1565	gtgaggtctgagacacttccctcaacaagaagc---gacctacagatgacagacgttgtgga	1621
D	b	1583	gtatgccaagcactactacactcaacgcgtctcagatgaaagttaactgtctgtccgcgcg	1642
Q	y	1622	aagaagaagtggcaactgtgaaggaagccagacactgtctccgcgcacgtgvtgttttgg	1681
D	b	1643	gttccatgtgtctctcaagaagaacatcacatctgtctccagcggcgccctacactctcgt	1702
Q	y	1682	ctttgctgtagcaacactcttgggtgtctgtctgacgtacaaacgctgtctgcgtcgcgc	1741
D	b	1703	agvtggccagaaagtgcaccaacatcgtgtgtggccaatactgtgcgtccctggtctttcag	1762
Q	y	1742	tgtcttggaactgtcgtgcctgtgttgcctgtggcacctaaagacccctgttgggtgaagtca	1801
D	b	1763	cactgccaattctttcatctctctgtgagaatcttccagacaccatgtgtgcgtccgcg	1822
Q	y	1802	ggggcgccgtgtcttctctatgtcgtgcctcctgtgcagcaggtatgtgcagctctatg	1861
D	b	1823	gtggccccaatgtgtcttctctgtatcgtgcgtccctgtcgtctgtgtgtgtgagtggccg	1882
Q	y	1862	gctctctttgggaaccccaagccagcgtgcgtgtgtgtgaagcagggcccttgtgcctg	1921
D	b	1883	tgtatgtggggccccccaacgctctctctcatgtctctgcgcgaagggcttcttcaacgt	1942
Q	y	1922	gtttccacatctctcgtctcgtcgtcgtacagttcgtcattccacaaatcatcatctca	1961
D	b	1943	gtcttccatctctctactcctctgataccgtgtgcgtctctccatacgaatcgtgtgtct	2002
Q	y	1982	agtttccacccaaggtacccatacttaccacgctcgtgtccaacaaacagcgtgtgtgc	2041
D	b	2003	agatgtgcgaagcgcgtgcgaagtgtcctaaagttttgtgtgtgtgtatccaagggccct	2062
Q	y	2042	tgtttgtgatgatcagctcagcggcccaagctg---cttatctgtctaaactgtgcgtgtg	2098
D	b	2063	tcttcgtgctctatcatcagcgcataaggtgtgcctgtgtgtggcaaatcatgtcgtgca	2122
Q	y	2099	tgtgtgaccccaactgcgtcgtcagggaaatataccagcgtcttccccatctgtgtgatgt	2158
D	b	2123	ccaccatcaaaccccatctgtgcgggacccgcgaatgtaccccaacatactgactctcgt	2182
Q	y	2159	gcacagagacaaactcctctgtgcttcatatactgtgccttccctacaaatgagccctctca	2218
D	b	2183	gcacccttaactaacgaaagggctactgtttcaacacagcatgtgaacttgtcgtcgtctg	2242
Q	y	2219	tcaatgtccttgcctctgacgtacatcgtgtgaagaaactgtccagagaaactacaagagca	2278
D	b	2243	tgtcgtgtttcgaactctgcgtctatactgtgcaagagagctgcgccacaactacaagagca	2302
Q	y	2279	aatgtgtcaacttcaagcgtcgtcttcaaatctgtgtgcctgagctgttcaaccagag	2338
D	b	2303	agtttcaacacttcaagatgaactcttctctcaacccctccactctcctctgcacttca	2362
Q	y	2339	ccagcgtctaaagcagcaagtacatcgtcgtgcggccaacaatgaatgtgtgcgtgaagcag	2398
D	b	2363	tgtctgtgcaagcagcgtgtcgtgtacatcatatgtgaactcctgttcaactgttcaact	2422
Q	y	2399	tgaagcagcgtcctcgtgtgtgtatttctcgtctaagtgtacgtgtgaactcctgtgcgcag	2456
D	b	2423	tctcgtgcacatcgtccttggatctacttggccccaagtgtatactgacttcttcaaccg	2482
Q	y	2459	acctcaacacacagagacacttccagcgtctccatcattcagactacaacggcgog	2510
D	b	2483	agcgcaacactcaagcctatttcaatagatgtatccaaaggtgtaccacatvtga	2534

RESULT 15
AAZ58964

ID	AAZ58964 standard; DNA; 2532 BP.
XX	
AC	AAZ58964;
XX	
DT	08-MAY-2000 (first entry)
XX	
DE	Mouse GPCR-B4 polypeptide encoding DNA.
XX	
KW	Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; mouse;
KW	taste transduction pathway; taste receptor; foliate; fungiform; food;
KW	circumvallate; taste signaling; pharmaceutical; ds.
XX	
OS	Mus sp.
FH	
FT	Key Location/Qualifiers
FT	CDS 1..2532
FT	/tag= "a"
FT	/product= "GPCR-B4"
XX	
PN	MO200006593-A1.
XX	
PD	10-FEB-2000.
XX	
PF	27-JUL-1999; 99WO-US17104.
PR	
PR	28-JUL-1998; 98US-0095464.
PR	17-Dec-1998; 98US-0112747.
XX	
PA	(REBC) UNIV CALIFORNIA.
XX	
P1	Zuker CS, Adler JE, Lindemeier J;
DR	WPI; 2000-195257/17.
DR	P-PDB; AAT77357.
PT	New isolated sensory transduction G-protein coupled receptor, useful
PT	for developing products for use in studying and modulating the taste
PT	transduction pathway and for generating taste topographic maps -
XX	
PS	Claim 5; Page 71-72; 76pp; English.
XX	
CC	The invention provides nucleic acids encoding rat, mouse and human
CC	sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
CC	GPCR polypeptides are components of the taste transduction pathway. The
CC	nucleic acids can be used to identify taste cells and as tools for the
CC	generation of taste topographic maps that elucidate the relationship
CC	between the taste cells of the tongue and taste sensory neurons leading
CC	to taste centers in the brain. GPCR-B4 is useful as a nucleic acid probe
CC	for identifying subpopulations of taste receptor cells such as foliate,
CC	fungiform, and circumvallate taste receptor cells. The polypeptides can
CC	be used for identifying compounds that modulate sensory signaling in
CC	sensory cells. Such modulators of taste transduction are useful for
CC	pharmacological and genetic modulation of taste signaling pathways. These
CC	modulatory compounds can then be used in the food and pharmaceutical
CC	industries to customize taste. The present sequence represents a DNA
CC	encoding a mouse GPCR-B4 polypeptide.
XX	
SQ	Sequence 2532 BP; 498 A; 824 C; 631 G; 579 T; 0 other:
XX	

Query Match 16.6%; Score 420; DB 21; Length 2532;
 Best Local Similarity 50.6%; Pred. No. 4.9e-96;
 Matches 1241; Conservative 2; Mismatches 1167; Indels 42; Gaps 8

Dn	83 ccgagcttaccaccccgagattaccttcctgcagagcgtttcccttcattctgct 148
Oy	ctgacttaccaccccccggagatcttccttcgcagagcgtttcccttcattctgct 148
Dn	149 gtcgcagtggaagacaaaccgagggtgaccctgtgca--caggtctgtgatctta 205
Oy	gtcgcagtggaagacaaaccgagggtgaccctgtgca--caggtctgtgatctta 205
Dn	143 tgaagagtgtctcttaaccctcagctactctgcaggtgtgccaaagtgaatgaataacaactg 202
Oy	atgagatgtgtctaacacactcttccagctagtcatgtcggtgtgggtgtgaagaataaacaact 265

Db	203	agrgtctgggctcaacaactcaatgacagagccatgcatctgcgctggagagaaatcaacact	262
Qy	266	ccaagcgccctctgcccacaatcatcacccctggggtaacagctgatatgtgtgtgtgact	325
Db	263	gtatgctcttctgctgcgcggcggtgctgtctgcgtctagagatgtagtgtctgcactctc	322
Qy	326	ctgcacaatgtgtatgccaacgctgtgagatgctctccctgcgcgaagcgacaacacatlagc	385
Db	323	ccaacaatcatccacgctggtcctactctctctctgtcaacagatag--atgattcctgcacca	379
Qy	386	tccaagagagcccttcacactatctccctcaagtgctgcgacgtatgtgtgcgtctacagca	445
Qy	446	ccaacgctgtgcgccaaccaacagccgcctctgtaggccttcctctgtgtgcacatgtatgct	505
Db	440	ctgagctctgcacatcacacggtgtccacaatctctctcactctccctgcgtgcacaagttcacat	499
Qy	506	atgagcgccaaacagcgaagacgtctcaagcgtgtgaaagccgagatccctctctctgcgcaca	565
Db	500	atagcgcacatccacgcgaacagctgtcaagaaagccggtcctccctgcacatgctgcacatg	559
Qy	566	tcaccaatgacaagtaccaggtgtgagacaatgtgctgcgtcgtgcagagaagtctgggtgga	625
Db	560	tgcacagcgccacacccacacatcagagcgcattgtgcaactgtgtgttcacttccagtgga	619
Qy	626	ctgtgattctctctgtgtgtgcagcagatgagaaactatgggc--agctaaagggtgcagcac	682
Db	620	actgtgaactcgtgtgtgtgtgtgtagcgtgaatgaaatattatgtgcgaagagaaacagccactgtcga	679
Qy	683	tggagaaactcagcgacactgtgtcaaggtatctgattgtgttccaagagacatcatgcctctc	742
Db	680	gcacagcgtctcgacacaacactcggcgaatactgtcatgtgcctccacgagagtcttgcgccgtac	739
Qy	743	ctgcaccaagtgtggcgtatgtgaag-----atgcagtgtcctcatbgcgc	784
Db	740	cagaaaccaacacagcgtgtgtgagcgtctgtagagcagagacaacatgtgaacaactctcgtgaca	799
Qy	785	acctggcaccaagcgccagggccacacgctcgtgtgtgttttttccaagccggcagattggtccaggg	844
Db	800	agctcgtcgagaaactcgtgcgctgtgtgtgtatattctccgcggaaagctctgaagcctgcaca	859
Qy	845	tgttttctgcagtcctcgtgtgtgtatgccaaactctacatcgtgcagaagtgtgtgtgtccctccagaag	904
Db	860	actctctcgtgtggcgtgtgcgtcgtgaaacttcaacgggtcttgtgtgtgtatgtcctctgagt	919
Qy	905	ctgtgacccctctccaaagcacaatcatcgtgggtgtgcgggtatcccaagcgtacattggaatgtgc	964
Db	920	ctgtggtccatctgaacccgtgtcttacaacaactcatcaagagcgtgtggccacaagcgacacttcc	979
Qy	965	tgggcgtgtgcacatccacagaagaagggctgtccctcgtgcctgtgaagcgttttaagaagaactatg	1024
Db	980	tgggtgtgtacacatccacagaaggtgttccatctccctgtgtcttccagcgaacttccgaatgtgcacatg	1039
Qy	1025	ccggcgacagaacaagggcccccttgaagcctctgtgcacaagggcctcctgtgtgcagcagacatc	1084
Db	1040	acaagacca-----gggtatcgtcagyctccaacgagaaccagcctcgtgcggtactaccgttaacc	1093
Qy	1085	agctctgtacagaatacgaagaacttttatgtgtacacaacgatatgcacagatctcaagaactctct	1144
Db	1094	agagactcgtgcgcgtcgtatgaatcatcatcactgagttccttcaacaagttctcatgtcttcgg	1153
Qy	1145	ccatgagattctgcctacacaagcataccggtcgtgtgtatgtgcgtgtgcccatagtcctcaacc	1204
Db	1154	ggagcgtgtgtgtgtctacaagcgtgttaactcggcgtcttacaagcgtgtggccacacaccttccaca	1213
Qy	1205	agctcccggtcgttgcctctgtgcgtgtgttccaagggccggagtctaacccctggagacttt	1264
Db	1214	gacctccctccactgtcaatcaggtctcgtctgcacacaaggaatacgtctcatccaatgcagcgttac	1273
Qy	1265	tggagcagatccacaaggtgtcatcttctcttacaagaagacactgtgtgcgttaatatgaca	1324

Dd	1274	tcagggagatctgcgcatgtcaacttcaagctctctgvgcaacacgctctcttcttcgacgaac	1333
Qy	1325	acagagatcccttcagtagtactaatacatatgcccctggactggatggaatggaaccagaatgga	1384
Dd	1334	aaggaggacatgcagatgtctctcgagacatcatccagttgagcagtgvgggtcgtgacgaagac	1393
Qy	1385	cccttaacgctcgtgttccctccacatggtctccagttccagttacaacataatgaaccca	1444
Dd	1394	ccctccaaagacatcgctcctctactcccccacacggaagcaggtctgacacataatt--agca	1450
Qy	1445	aaatocagtgagcaggaaggaacgaacaggtgctcctaagtctgtgtgttccagcagctctc	1504
Dd	1451	atgtgtccctgggtacaccccccaacaacagtgctcccatatcatgtgttctaagagtgc	1510
Qy	1505	ttgaaggagccacacagcagatggtgttaagggtttccatcacctgcgtcttgagtgtgacct	1564
Dd	1511	agccctgggcaaatgtaaaaaaacccaataggctctccacccaatgctcttgagtggtgtgacct	1570
Qy	1565	gtcgggcttggagacctctccctccaacaagaagt--gaacctacagatgccaagcagcttvgga	1621
Dd	1571	gtccgcgcggaacacctaactcaacgcgatcagtagatgagtttaactgtctgtccctccgg	1630
Qy	1622	aagaagaatgvgcaccttgaggggaagccagacctgtcttccgcgcacgtgtgtgtttttg	1681
Dd	1631	gtccatcatgtgtcttacaagaacaacatcgtctgttcttaacggcgtgcgtccttcctg	1690
Qy	1682	cttttcgvtgagacacctctcttgggtgcgtgcgtgacagctaaacgctgtctgcgtgcgc	1741
Dd	1691	agtgagcagcaagtgcaccaatcatcgtgtgagaccaacctgcgcgcgcctctgagttcaagta	1750
Qy	1742	tgcctgggagctctctgacctgtttgtctgtctgcaactaagaaaccccttvggtgagatcaag	1801
Dd	1751	cgctggccaatctgtctcaatcttctcggagacattccacagacgcgaatggtcgtctcgsg	1810
Qy	1802	ggggccgcgctgtgtcttctttaatgctcgtgggtccctcgtgacagcaggtatgtgcagctctatg	1861
Dd	1811	ggcgccccaatggtcttccctgaatgcgtgtgcgtgcctctgtcgtgcgtctcggatgtgccccg	1870
Qy	1862	gcttctcttgggaacccaagaagcctgctgctgtgtctgaagccagggccctcttgcctctg	1921
Dd	1871	tgtatgtgggcccccccacaggtctctcctgttctctgtctgcgcgcaggtcttctcaacgttt	1930
Qy	1922	gtttcaacatctctctgtctcgtctgacagatgtgcctatcccaataatcatcatcatct	1981
Dd	1931	gctctcgcgtctgcctctctctcatcagatgctgcctctccagattgtgtgtgtcttca	1990
Qy	1982	agtttccaacgaagtactaatctcacaacgcctggtgtccaacaacaggtgtctggc	2041
Dd	1991	agatgbcagcaagcctgcgaacagcgcctacggttcttctgagtgcgttacaacagggccctacg	2050
Qy	2042	tgttggtagtgaagcttcagcggccagcgtg---ctatcgtctaatctgscgtggtg	2098
Dd	2051	tctctgtgctcttcaacagggccgttcaagggtggtccctgtgtgctgggcaacatgctgtgcca	2110
Qy	2099	tgttgaaacccaatgtgcctctaggaagataacacagcgtcttccccaatctgtgtgactgtgagt	2158
Dd	2111	ccacatcaaacccatgttggcggagacggaaccccgatgagcccaatatcatatccctctct	2170
Qy	2159	gcacagagaccaaactccctgtggcttcatatctggtcctctcctbacaatgagctcctccca	2218
Dd	2171	gcacccctaactacgcgaacgggtgctactcttcaacaacagcagtagaactgtgtgtctcg	2230
Qy	2219	tcaagtccttgcctgcgaagctacacgtgttaagagcttccagagatacagaagcaggtgca	2278
Dd	2231	tgtctggttcaagctctgcgtacgtggtggaagaaagctgcaccaacatacaacgaagacca	2290
Qy	2279	aatgvtcaactccacgcgtcctctcaactctgttctctgtagtgcctctcttccacaacgg	2338
Dd	2291	agttcatcaacctcaagatgaccttctccttaactctctcaatctccctctgtgcaagttca	2350
Qy	2339	ccagcgtctacagcgcaagtacactcgtcctgctggccaacatgagtgtgcgtgtgagcagcc	2398
Dd	2351	tgtctctcaacatgagctgcgtcgtgcacacatcatgactctcctgtgtcaactgtgtcactat	2410

QY 2399 tgaagcagcggtcgcgtgggtatttctgcctaagtgtaogtgaatccctgcgcgcag 2458
Db 2411 ttctggccatcggtctgggtacttggccccaatgttaacatgaccccttctaccg 2470
QY 2459 acctcaacagacagacacttccaggcctccattcagactaacagagcg 2510
Db 2471 agcgcaacacttcagcttattcattagcatgaggtacacagatgag 2522

Search completed: June 11, 2002, 06:48:12
Job time: 10293 sec

QY 2101 tgaacccactgctctgtaaggaaataaccagcgtctcccccactctgctgtagcttgatgac 2160
|||||
Db 2101 TGAACCCACTGCTCTGTAGGAAATACAGCGCTTCCCCCATCTGTGTATCTGTAGTGC 2160
QY 2161 acagagacaaactccctggtctatactatgctctcctctacaatgagcctctcctc 2220
|||||
Db 2161 ACAGAGAACAACTCCCTGGGCTTACTGTGCTCTCTCAATGAGCCTCTCTCCATC 2220
QY 2221 agtgcctctgctcagctagctatgctgtaaggacttgccaggaactaaagaagccaa 2280
|||||
Db 2221 AGTGCCTTGTGCTGAGTACTGTGGTAAGGACTTGGCCAGAGAACTAACAGAGGCCAAA 2280
QY 2281 tgytcaacctcagcctgctcttcaactctgtccctgagatcgccctcttcaacagagcc 2340
|||||
Db 2281 TGYTCAACCTTACAGCTGCTCTTCACTGCTGTCTGTGATCGCTTCTTCAACAGGCC 2340
QY 2341 agcgtctacagagcgaagtaactctgctgagccaacatgagtgtggtctgagagcctg 2400
|||||
Db 2341 AGCCTTACGACGCGCAAGTACTGCTGCGGCCAAACATGATGGCTGGGCTGAGAGCCTG 2400
QY 2401 agcagagctcggtggtgattcttcgtctaagtgtacgtagatccctctgcccagac 2460
|||||
Db 2401 AGCAGCGGCTTGGTGGGTATTTCTGCTAAGTCTACGTGATCTCTGCGGCCAGAGC 2460
QY 2461 ctcaacacagagagcactctcagagcctccatcaagagctacacagagcgctgcgctcc 2520
|||||
Db 2461 CTCAACACAGACAGACACTTCCAGGCCCTCCATTCAGAGCTACAGAGGCGCTCGCGCTCC 2520
QY 2521 acctga 2526
|||||
Db 2521 ACCTGA 2526

RESULT 2
AX282900
LOCUS AX282900 2951 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 9 from Patent WO0172842.
ACCESSION AX282900
VERSION AX282900.1 GI:16609872
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (sites)
AUTHORS Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,
Abuin,A., Zambrowicz,B. and Sands,A.T.
TITLE Novel human 7tm proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0172842-A 9 04-OCY-2001;
Lexicon Genetics Incorporated (US)

FEATURES
source
1..2951
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 559 a 888 c 831 g 669 t 4 others
ORIGIN

Query Match 100.0%; Score 2524.8; DB 6; Length 2951;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgtctctcaagcgtcgctgtgctgagctgagcttcattctctgtgtcggcc 60
|||||
Db 232 ATGTGCTCTCAAGCGCTCGCTGTGCTGAGCTTCATTTCTCTGCTGCGGCC 291
QY 61 ttctgcctcagtagaagagatctctcctgaactcaacccctccgagatatacctctg 120
|||||
Db 292 TTTCGCTGCCATGACAGCGAGTCTTCTCTACCTTCAACCTCCCGAGATATTACCTCTG 351
QY 121 gcagagctgtctcctctcattctgtgctgtctgcaagtgagacagaccgagtgacc 180
|||||
Db 352 GCAGGCGCTGTCCCTCTCCATTCTGCGGTCTGCGAGGTGAGACAGACCAGAGGTGACC 411

QY 181 ctgtctgacaggtctctgtagcttcaatgagcatgagctacacacctctccagagctatgcg 240
|||||
Db 412 CTGTGTACAGGTCTTGTAGCTTCAATGAGCATGGCTTACCACTTCTTCCAGGCTATGGCG 471
QY 241 ctgtggtctgagagataaacaactccacagccctctgcccacaatcaccttggtgac 300
|||||
Db 472 CTTGGGGTGTGAGAGATTAACACACTCCACGGCCCTGCTGCCAACATCACCTGGGGTAC 531
QY 301 cagctgtatgctgtgtgtgtgaactctgccaatgtgtatgtccacgctgagagtgctctcc 360
|||||
Db 532 CAGCTGTATGATGTGTGTGTGACTCTCCATGTGTGTGACAGCTGAGTGTGCTTCC 591
QY 361 ctgtccagagagacacccacatagagcttccaaagaagacctctccactatccctcaaggtg 420
|||||
Db 592 CTGCCAGGGCAACACACATGAGCTCCAAAGAGACCTTCTCCACTATTCCTCTACGGGTG 651
QY 421 ctgscagtgatctggtcctgacagacccaacgctgtctgcacacacagccgctctgagc 480
|||||
Db 652 CTGGGAGTATTTGGGCTTGACAGACCAACCGTGTGCTGCACACAGCCGCCCTGCTGAGC 711
QY 481 ccttctctgtgtgccatgatagtatgctgagccagcagcagagcgtcagctgaagcgg 540
|||||
Db 712 CTTTCTCTGTGCTCCCATATTAGCTATGCGGCCAGACGAGAGCTTCAGCGTGAACGG 771
QY 541 cagatatacctcttctctctgctgcagacatacccaatgatacaagtgtaacagatgtg 600
|||||
Db 772 CAGTATCCCTCTTCTCTGCGACCATCCCAATGACAAAGTGAAGTGAACCATGTGTG 831
QY 601 ctgctgtctgacgaagctcgtggtgagcctgatactctctctgtctgagcagtgacagactat 660
|||||
Db 832 CTGCTGCTGACAGAGTGTGGGTGAGACCTGATCTCTCTGTGCTGAGAGAGACACTAT 891
QY 661 gggcagctaaaggtgtcagagcaactgagaaacacagccactgtgtcagggatctgcaatgt 720
|||||
Db 892 GGGCAGCTAGGGGTGACAGGCACTGAGAACAGGCACTGCTCAGGGGATGTGCACTTGTCT 951
QY 721 ttcaagagcatcaagccctctctctgcccaggttgagagtgagagatgcaatgtctatgt 780
|||||
Db 952 TTCAAGGACATCATGCTCTTCTCTGCCCCAGGTGGGCGATGAGAGATGACATGCTCTCATG 1011
QY 781 cgcacactgtgcccacagccggggccacgcgtctgtgtgttttttccaagccggcagcttgacc 840
|||||
Db 1012 CGCACACTGGGCGCAGGCGGGGCCACACGTCGTGTTGTTTTCAGCGGCGCATGTGGCC 1071
QY 841 aggtgttttctcagagctcgtgtgtgtctgacacactgactgtgcaaggtgtgtgtgtgtc 900
|||||
Db 1072 AGGGTGTTTTTCAGAGTCCGTGTGTGACCAACCTGACTGCGCAAGGTGTGGTGCCTTCA 1131
QY 901 gaaacgtgtggccctctccagacacatactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
|||||
Db 1132 GAAGCTTGGGCTCTCTCCAGGCATCTACTGTGGGTGCCCCGGGATGCCAGCATTTGGATG 1191
QY 961 gtgtctgagctgtgcatccacaagaagaggtgtctcctgtgacgtgaagcgcttgaagaagcc 1020
|||||
Db 1192 GTGCTGGGCGTGGGCATTCAGAGAGGGCTGTCCCTGTGCTGAAGGCTTTGAAGAAGCC 1251
QY 1021 tatgtccgggacagacaagagagccctcagcctgtgccaagaggtctcgtgtgcaagc 1080
|||||
Db 1252 TATGCCCCGCGACAGACAMAGAGGCCCTTGTGCTTGCACAGAGGGCTCTGTGTGACGACG 1311
QY 1081 aatacagctctcagaagaatgtcgaagcttctatgtcaacacagatgtgtccaaactcaaac 1140
|||||
Db 1312 AATCAGCTCTCAGGAATGTCCAACTTTCATGTACACAGATGCCCCAAAGCTCAAAAGCC 1371
QY 1141 ttctccatgagtctctgcttacaacgcatataccgggtgtgtatgtcggtgtccatgtgctctc 1200
|||||
Db 1372 TTTCATGATGATTTGCTTACACGATACCGGGGTGTGTGCGGTGTGCCATAGGCTTC 1431
QY 1201 caccagctctgtggtgtgtgtctctgtgagactgtgtccaggggacaggtctacccctgtgag 1260
|||||
Db 1432 CACCAAGCTCTCGGGGTGTGCTGTGAGACTTGTTCAGGGGCGCAGTGTACCTCTGGCAG 1491

```

OY 1261 cttttagagacagatccacaagtgatcttctctacacaagacactgtgagctttaat 1320
    |||
DB 1492 ctttttagacagatccacaagtgatcttctctacacaagacactgtgagctttaat 1551
    |||
OY 1321 gacaaagagatccctctagtagctataacataatgtccttgagacacggaatgagcccaag 1380
    |||
DB 1552 gacaaagagatccctctagtagctataacataatgtccttgagacacggaatgagcccaag 1611
    |||
OY 1381 tgaacctcaagctcctctggttctccacatgtctccagatcagttacaataacataatgag 1440
    |||
DB 1612 tgaacctcaagctcctctggttctccacatgtctccagatcagttacaataacataatgag 1671
    |||
OY 1441 accaataatccagtaggacaggaacaaacaaagtgacctaaagtcgtgltgtgtccagagac 1500
    |||
DB 1672 accaataatccagtaggacaggaacaaacaaagtgacctaaagtcgtgltgtgtccagagac 1731
    |||
OY 1501 tttcttgaaggagacacagagagatgtgttaacggtttccatcactgctgtttgagtggtg 1560
    |||
DB 1732 tttcttgaaggagacacagagagatgtgttaacggtttccatcactgctgtttgagtggtg 1791
    |||
OY 1561 cccctgtgagctgtgagaccttctccacaagagtgacctatacagatgccaagcctgtgag 1620
    |||
DB 1792 cccctgtgagctgtgagaccttctccacaagagtgacctatacagatgccaagcctgtgag 1851
    |||
OY 1621 aaagaagaatgagacacacacacacacacacacacacacacacacacacacacacacacac 1680
    |||
DB 1852 aaagaagaatgagacacacacacacacacacacacacacacacacacacacacacacacac 1911
    |||
OY 1681 gcttgcgtgagacacacacacacacacacacacacacacacacacacacacacacacacac 1740
    |||
DB 1912 gcttgcgtgagacacacacacacacacacacacacacacacacacacacacacacacacac 1971
    |||
OY 1741 ctgtctgtgagacacacacacacacacacacacacacacacacacacacacacacacacac 1800
    |||
DB 1972 ctgtctgtgagacacacacacacacacacacacacacacacacacacacacacacacacac 2031
    |||
OY 1801 gggggcgccctgtgtcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1860
    |||
DB 2032 gggggcgccctgtgtcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2091
    |||
OY 1861 ggtcttcttggggaacacacacacacacacacacacacacacacacacacacacacacacac 1920
    |||
DB 2092 ggtcttcttggggaacacacacacacacacacacacacacacacacacacacacacacacac 2151
    |||
OY 1921 ggttccacatcttctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 1980
    |||
DB 2152 ggttccacatcttctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 2211
    |||
OY 1981 aagtttccacaagaatgacatacttctacacacacacacacacacacacacacacacacacac 2040
    |||
DB 2212 aagtttccacaagaatgacatacttctacacacacacacacacacacacacacacacacacac 2271
    |||
OY 2041 cgtgtgtgataatgacacacacacacacacacacacacacacacacacacacacacacacac 2100
    |||
DB 2272 cgtgtgtgataatgacacacacacacacacacacacacacacacacacacacacacacacac 2331
    |||
OY 2101 tggagcccaactgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2160
    |||
DB 2332 tggagcccaactgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2391
    |||
OY 2161 acaagaaacacacacacacacacacacacacacacacacacacacacacacacacacacac 2220
    |||
DB 2392 acaagaaacacacacacacacacacacacacacacacacacacacacacacacacacacac 2451
    |||
OY 2221 aatgtccttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2280
    |||
DB 2452 aatgtccttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2511
    |||
OY 2281 tgytgaacttcaagcctgtcttcaactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2340
    |||
DB 2512 tgytgaacttcaagcctgtcttcaactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2571
    |||
OY 2341 agcgcttacaagcgcaactgacacacacacacacacacacacacacacacacacacacacacac 2400
    |||

```

```

DB 2572 AGCGTCTACAGCGGCAAGTACCTGCTGGCGCCAAATATAGCTGAGCTGAGCACCTGG 2631
    |||
OY 2401 agcagcgcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2460
    |||
DB 2632 AGCAGCGGCTTGGTGGGATTTTCTGCTTAAGTGTCTAGTGTCTCTGCTGCGCCAGAC 2691
    |||
OY 2461 ctcaacagacagagacacttccagcctccatccatcaagactacaagagcgctgtgagctcc 2520
    |||
DB 2692 CTCACAGACAGACAGACATTCAGCGGCTTCATTCAGAGACTACACAGAGCGGCTGGCTCC 2751
    |||
OY 2521 acctga 2526
    |||
DB 2752 ACCTGA 2757
    |||

RESULT 3
AX354027
LOCUS AX354027 2759 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 13 from Patent WO0203848.
ACCESSION AX354027
VERSION AX354027.1 GI:18618964
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
AUTHORS Naito A.T.
TITLE Method for opening the blood-brain barrier
JOURNAL Patent: WO 0203848-A 13 17-JAN-2002;
Naito, Albert T. (US)
FEATURES
source 1..2759
location/Qualifiers
BASE COUNT 523 a 828 c 782 g 626 t
ORIGIN

Query Match 99.98% Score 2523.2; DB 6; Length 2759;
Best Local Similarity 99.88% Pred. No. 0;
Matches 2522; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 atgctgtctgcaagcgtcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 60
    |||
DB 53 ATGCTGCTGTGACAGGCTGCGCTGGTGGCTGACGCTTCTCATTTCTGCTGCTGGGCC 112
    |||
OY 61 ttgtctgcatagacaggaagttcttctgacttcaacctcccgagattactctctg 120
    |||
DB 113 TTGGCTGCCATAGACAGGAGTCTTCTGACTTCAACCTCCCGAGATTAACCTCTG 172
    |||
OY 121 gaagcgcttccctccatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
    |||
DB 173 GCAGGCGCTTCCCTCTCATTTCTGCTGTGTGACGTAGGACACAGACCGAGGTCGACC 232
    |||
OY 181 cgtgtgacagctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
    |||
DB 233 CTGTGTGACAGGCTTGTGTGCTTCAATGACATGGCTTACACCTCTTCCAGGCTATGCGG 292
    |||
OY 241 ctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
    |||
DB 293 CTTGGGGTTGAGAGATTAACAACACTCCAGGCGCTGCTGCCAACAATCAACCTGGGGTAC 352
    |||
OY 301 cagctgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360
    |||
DB 353 CAGCTGTATGATGTGTGTCTGACTCTGCCAATGTGTATGCCACGCTGAGTGTGCTTCC 412
    |||
OY 361 ctgtcagggcaacacacacacacacacacacacacacacacacacacacacacacacacacacac 420
    |||
DB 413 CTGCGAGGGCAACCACTATGAGCTCCAAAGGAGACTTCTCACTATTCCCTACGCGTGG 472
    |||

```


AX282894
LOCUS AX282894 2292 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 3 from Patent WO0172842.
ACCESSION AX282894
VERSION AX282894.1 GI:16609869
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,
Abdin,A., Zambrowicz,B. and Sands,A.T.
TITLE Novel human 7tm proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0172842-A 3 04-OCT-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source 1..2292
location/Qualifiers
BASE COUNT 437 a 690 c 638 g 524 t 3 others
ORIGIN
Query Match 90.7%; Score 2290.8; DB 6; Length 2292;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 235 atgagcggttgaggttgagagataaacaactcacagcgccctctccacacacacccctg 294
DB 1 ATGCGGCTTGAGGCTGAGAGATAAACAACACTCACAGCGCCCTCTGCCAATCACCTCG 60
QY 295 gggtaacagcgtlatagatgt 354
DB 61 GGGTACAGAGCTGTATGATGT 120
QY 355 ctctccctgcagaggaac 414
DB 121 CTTCT 180
QY 415 acggt 474
DB 181 ACGGT 240
QY 475 ctgagcccttctctgt 534
DB 241 CTGAGCCCTTCTCTGT 300
QY 535 aagcgagcagatccct 594
DB 301 AAGCGGCAATATCCCTTCTCTGT 360
QY 595 atggt 654
DB 361 ATGT 420
QY 655 gactatgagcagctagaggt 714
DB 421 GACTATGAGCAGCTAGAGGT 480
QY 715 atgtgttttaagaagacataatgtccctctctctctctctctctctctctctctctctct 774
DB 481 ATTTGCTTTTAAAGACATCATGCT 540
QY 775 ctcatgagccacactgagccagagcgagcgacgctgtgtgtgtgtgtgtgtgtgtgtgtgt 834
DB 541 CTCATGCGCGACACTGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 835 ttggcagaggt 894
DB 601 TTGGCAGAGGT 660
QY 895 gctctcaagaagccttgcgacctctccagagcacaatcactgtgtgtgtgtgtgtgtgtgt 954

DB 661 GCCTAGAGAGCCTGTGGGCT 720
QY 955 gggatgt 1014
DB 721 GGGATGT 780
QY 1015 gaagcctatgcccgggagacagaagagggcccttaggccttgcacaaagggtcctgtgtgc 1074
DB 781 GAAGCCTATGCCCGGGAGACAGAGAGGGCCCTAGGCGCTTGGCAAGAGGGCTCTGTGTGC 840
QY 1075 aagcagaatcagactctgtcagagaatgtccaaagcttctatgtcacacaagatgtccaaagctc 1134
DB 841 AACAGCAATACACTCTGTGAGAAATGCCAAGGTTTATATGACACACAGATGCGCAAGCTC 900
QY 1135 aaagccttccatgtatgtctgtcctcaacagacataccgggctgtgtatgcgtgtggccat 1194
DB 901 AAGCCTTCTCTATGATGT 960
QY 1195 ggcctccacagactcctgt 1254
DB 961 GGCCTCCACAGCTCCTGTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 1020
QY 1255 tggcagcttttggagcagatcccaaggtgtatcttccctctacaaaggacactgtgtgc 1314
DB 1021 TGGCAGCTTTTGGAGCAGATCCACAGAGGTGATTTCTCTACACAGAGACACTGTGGCG 1080
QY 1315 ttaatgacaacagagatccctcagtagctataacataatgtgtgtgtgtgtgtgtgtgtgt 1374
DB 1081 TTTAATGACAAAGAGATCCCTCAGTAGCTATTAATATATGCTCTGGAGCTGGAATGGA 1140
QY 1375 cccaagtgtacatcaagctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1434
DB 1141 CCCAAGTGTACCTTACAGCTTCTCGTGTCTCTACATGTGTCTCAAGTTCCAGTTAAACATA 1200
QY 1435 aatgagacaaaatccaggt 1494
DB 1201 AATGAGACCAAAATCCAGT 1260
QY 1495 agcagactgtctgtgaagggagcagcagcagcagcagcagcagcagcagcagcagcagcag 1554
DB 1261 ACGCAGTGTCTTGAAGGGAGCCAGCAGAGGTGTACGGGTTCATACACAGCTGCTTTAG 1320
QY 1555 tgt 1614
DB 1321 TGT 1380
QY 1615 tgt 1674
DB 1381 TGT 1440
QY 1675 ttttggcttgcgt 1734
DB 1441 TTTTGGCTTTCGT 1500
QY 1735 ctgt 1794
DB 1501 CTGT 1560
QY 1795 tcaagcaggggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1854
DB 1561 TCAAGCAGGGGGGCGCTGT 1620
QY 1855 ctctatgtcttcttggggaacccaagagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1914
DB 1621 CTTATGCTTCTTGGGGAACCAAGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
QY 1915 ggccttgttttcaacatctcctgtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1974
DB 1681 GCCCTTGT 1740
QY 1975 atcttcaagttttcaacaaagttacttacttacttacttacttacttacttacttacttact 2034

Db 1741 ATCTTCAAGTTTTCACCAAGTACTACTATTTACACAGCCTGGTCCAAACACAGGT 1800
 QY 2035 gctgacctgtttgatgatcagctcagcgccagctgtctatcttctactgtctg 2094
 Db 1801 GCTGGCCGTTTGTGATGATAGCTCAGCGGCCACGCTGCTTATCTGCTACTTGGCTG 1860
 QY 2095 gtgtgtgtgagcccccacgtcgtcctcaggaataacacagcgtctcccccattctgtatgctt 2154
 Db 1861 GTGTTGTGGAGACCCACAGCTGCGTGTAGGGAATACAGAGCGCTTCCGCCATCTGTGATGCTT 1920
 QY 2155 gaatgcacagagacacacactccctgtggtcttaactgtgctctctctactatgtcctc 2214
 Db 1921 GAGTGCACAGAGACCAACTCCCTGGGCTTACTGTGGCTTCTCTACAAATGGCTCCCTC 1980
 QY 2215 tccatagtgcttctgtctgagactcgtgtgtaagactgtgcagagaaataacacag 2274
 Db 1981 TCCATGAGTGGCTTGTGCTGAGCTACTGTGGTAAGACTTGCCAGAGAACTTACAAACAG 2040
 QY 2275 gccaaatgtgtcaccttcaagctgtccttcaactcgtgtctcgtgctgctctctcacc 2334
 Db 2041 GCCAAATGTGTACACTTCAAGCTGCTGTTCAACTTCGTTCTGTGATGCGCTTCTTCAAC 2100
 QY 2335 acggccacagctctacagacgaagtaacctcgtcgtcgcccaacatgatgtgctggtgac 2394
 Db 2101 ACGGCCAGCGCTTACGACGCGCAAGTACTGCTGCGGCCAACATGATGCTGGGCTGAGC 2160
 QY 2395 agcttgagcagcgcttcgtggtggtatttctgcttaagtgtctgagatcctctcgtcgcg 2454
 Db 2161 AGCTGACACAGCGCTTGGGGGATTTTCTGCTTAAGTCTGATCGTATCCTTGGCGGC 2220
 QY 2455 ccagacacctcaacacagacagacacttccagcgctccattcagactacacagacagcgctgc 2514
 Db 2221 CCAGACCTCAACAGACAGACGACTTCCAGGCGCTCCATTCAGAGATACAGAGGCGTGC 2280
 QY 2515 ggtccacactga 2526
 Db 2281 GGCCTCACCTGA 2292
 RESULT 5
 LOCUS AF301162 2526 bp mRNA linear ROD 12-FEB-2001
 DEFINITION Mus musculus strain 129P3/J putative sweet taste receptor T1R1
 ACCESSION AF301162
 VERSION AF301162.1 GI:12745519
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2526)
 L1.X., Inoue,M., Reed,D.R., Hugue,T., Puchalski,R.B., Tordoff,M.G., Nishimura,Y., Beauchamp,G.K. and Bachmanov,A.A.
 High-resolution genetic mapping of the saccharin preference locus (sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to mouse distal Chromosome 4
 Mamm. Genome 12 (1), 13-16 (2001)
 JOURNAL MEDLINE 21030739
 REFERENCE 2 (bases 1 to 2526)
 L1.X., Reed,D.R., Hugue,T., Puchalski,R.B., Tordoff,M.G., Beauchamp,G.K. and Bachmanov,A.A.
 Direct Submission
 Submitted (29-AUG-2000) Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104-3308, USA
 FEATURES
 source
 1. .2526
 /organism="Mus musculus"
 /strain="129P3/J"
 /db_xref="taxon:10090"
 /chromosome="4"
 /map="Between D4Mit42 and D4Mit254"
 1. .>2526
 gene

CDS
 /gene="Gpr70"
 1. .>2526
 /gene="Gpr70"
 /codon_start=1
 /product="putative sweet taste receptor T1R1"
 /protein_id="AAK07092.1"
 /db_xref="GI:12745520"
 /translation="MLPMAHLLLSLQLAAYCAWAFSCORTESPSPGSLPDELGLAGL
 FSLHADTQVHRPIVYSCDRSDSPSGHVLHROAMFPEYVEELINSPALLPNTTLYGE
 LHYDCESSNMYATLRYLAQGGHLEMDQDLNRHSSKVALLGPDNTHAVTAAAL
 SPFLMPLVSTEASSVILSGKRKPFSELTARTIPSDKYDEVIVRLQSGRWVLSVGSY
 GDYGOVLQALELAPRGICVAEKDVPVLSAAGDPRMGRMLRLARATTVVVS
 NRHLGAFPRSVYLANITGKVMVLAISEDMALISTYITNPGIIGTGVAVIQORVPG
 LKEFEESYVOAVTGAPRTCEGSGNCSNARECHAFATTNMPELGFASMAVNYE
 AYAAYAHGLHQLIGTSSTCARGPVYPMOOLQIKYNFLHKKTYAFDNDGPIGY
 DILAMPNGSEPMTEVYIGSASLSPVHDIKRTIQMHKKNQVPSVCTDCEBHR
 LVMSHNCPECEHCEAGTFLNLTSELTQPCPGDEAPRGSSCSRYEFLGMBEP
 ISLVLANTLILLLLIGTGLPAMRILHTVNSAGRLCFMLGSLVAGSCSLYSFE
 GEPTVPACILRQPLFSIGFAIFLSCLTIRSFQVLIIFKPESTKPTFEHTMAOHNGAI
 FVIVSTVHLECLTLMAMTPTPTREYORFPHVILECTEVSNGFLVAFANILLS
 ISTFVCSYLKELPENYNAKCVTFESLLHFVSIADFETMSIYGSYLPAVNVLAGL
 ATLSGSGSFLKRCYIILCRPELNTENHQAISIDYTRCGTT"
 BASE COUNT 483 a 760 c 679 g 604 t
 ORIGIN
 Query Match 67.0%; Score 1692.8; DB 10; Length 2526;
 Best Local Similarity 79.7%; Pred. No. 0;
 Matches 1994; Conservative 3; Mismatches 505; Indels 0; Gaps 0;
 QY 22 ctgttcgagcgcagcttcattcctcgtcgtcgtggtccttgcctcatalagacagag 81
 Db 25 CTGCTCAGACCGTCGACGCTGCGGTGCTTACTGCTGGGCTTTCAGGTGCCAAGAGACAGAA 84
 QY 82 tctctcctgaactcaacctcccgagataactcctctgcaagcgcttccctccat 141
 Db 85 TCCCTCCAGCTTTCAGGCTCCGCGGACTCTCCCTCGGAGCGCTGTTCTCCCTCAT 144
 QY 142 tctgagctgtcagagtgtaggcacagaccgagtgtaaccctgtgtgacagctctgtgac 201
 Db 145 GCTGACTTCTGTCAGGTAGACACAGACCTCTGGTGACAAATTGTGACAGGTCTGACAGC 204
 QY 202 ttcaatgacatctgctacacaccttccaggctacgtggtgtggtgtgagagataaac 261
 Db 205 TTCAACGGCCATNGGCTATCACCTCTTCCAGACGATNGCGTTACACGCTTGAGAGATAAAC 264
 QY 262 aactccacgcccgtcgtcccaacaatacctctggtgtacacagctgtatgtgtgt 321
 Db 265 AACTCCACAGCTGCTCTCCACATCACCTCGGGGTATGAACGTGTATGACGTGTCTCA 324
 QY 322 gactctgcacatgtgtatgcacagctgtaggtgtctctccctgcagggcaacacacata 381
 Db 325 GAGTCTTCCAAATGCTTATGCCACCTTGAGGGTGTCTCGCCCGCAAGGACAGCCACCTA 384
 QY 382 gactccaaagagacaccttccactatccctcactcagtgctgtcgtgacagtgattgggctgac 441
 Db 385 GAGATGCAGAGAGATCTTCCGCAACACACCTCCCAAGGAGGTGGGACATATTGGGCTGAT 444
 QY 442 agcaccaacgctgtgcacacacagcgccctgtgagccttctcctgtgtgccaatgatt 501
 Db 445 AACACTGACACGCGTGTCCACACTGCTGCTGAGACCCCTTTTCTATGATCCCTTGCTC 504
 QY 502 agctatgtgcacagcagagagagctgtagcgctgtagagggcaggtatccctcttctgagc 561
 Db 505 AGCTACGAGGAGAGAGGCTATCTCAGTGGAGAGGCCAAGTTCCCTCTCTTGGGC 564
 QY 562 accatcccaatgacaaataccagtgtagagacacatgtgtcgtcgtcgtcagaaatctggg 621
 Db 565 ACCATCCCGACGAGTAATACAGGTGAGATGATAGCGGCTGCTCAGACGCTTCCGC 624
 QY 622 tggactgtgactctctgtgttgagcagtgtagcactatgtggcagctaggggtgacgca 681
 Db 625 TGGCTGTGATCTGCTGTTGGGACGATGATGATGATGATGATGATGATGATGATGATGATG 684

QY 682 ctggaacaccagccactggtcagggatctgattgtcttcaaggacatcatccctc 741
 DB 685 CTGAGAGAGCTGGCAOCCACAGGGGATCTGCTGCTTCAAGAGCGTGGCTCTC 744
 QY 742 tctgccaggttggcgatgagagatgcatgctcatgcccactggccagccgg 801
 DB 745 TCCGCCAGCGGGGATCCCAAGATGCAAGCGCATGATGCTGCTGCTGCAAGCAG 804
 QY 802 gcaacgctggtgtgtttttccagcgagcttggccaggtgttttcagctcgt 861
 DB 805 ACCACCGTGTGCTGCTGCTTCTTAACCGGACCTGGCTGGAGATGTTCTCAAGTCTG 864
 QY 862 gtagtaccacactgactgcaagagtggtgctcctcagaagcctgggcccctcag 921
 DB 865 GTGCTGGCAACTGACTGGCAAGGTGGATGCGCTCGAAGACATGGGCACTCTCACG 924
 QY 922 caatactcgtgggtgctccgggataccagcattggatgtgtgtgtgtgtgtgtgt 981
 DB 925 TACATCACCAATGTGCTCCGGATCCAGGGATTTGGACGGTGTGGGGGTGGCCATCCAG 984
 QY 982 aagagggctgctccctgagcgtgttgaagaagcctatgcccgggagagaagag 1041
 DB 985 CAGAGACAGTCCCTGCTGCTGGAAGAGTTTGAAGAGTCTATGTCAGAGCATGACGGGT 1044
 QY 1042 gcccttaggcttgcacacaagggctcgtgtgagcagaacatcagctctgcaagatgc 1101
 DB 1045 GCTCCAGAACTTGCCCAAGAGGGGTCTGCTGGCGCATTAACAGCTGTGCGGAGTGT 1104
 QY 1102 caagcttatgtacacacagatgcccagcttcaagccttccatgtgtgtgtgtgtgt 1161
 DB 1105 CAGGCTTTACGACATGAAACATGCCCGAGCTGGAGCTTCTCCATGAGCGGTGCTTAC 1164
 QY 1162 aacgcataccaggtgtgtatgctggtgagccatgctcaccagctcctggtgtgtgt 1221
 DB 1165 AATGTGTAGAGAGCTGTGATGTGTGTGGCCACAGGCTCTCACAGCTCTCGGATGTACC 1224
 QY 1222 tctgagctgtgtcaggggagcagctcaccctgagcatttggagaagatccaaag 1281
 DB 1225 TCTGGAGACTGTGCGAGAGGCCAGCTACCCCTGCGACCTTCTCAGAGATCTACAG 1284
 QY 1282 gtagcatcttcttacaagaagcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1341
 DB 1285 GTGAATTTCTTCTTACATTAAGAGACTGTGACTTGTATGACACAGCGGACCTCTAGGT 1344
 QY 1342 agctataacataatgtcctggtgactgtgaatggaacccaagtgagcctcagctcctgt 1401
 DB 1345 TACTATGACATCATGCTGCGAGCTGGAGTGGGCTGAATGAGACCTTTGAGGTCTTGGT 1404
 QY 1402 tctccacatgtgtcctcagcttcaagcctaaacataatgagaacaaaatccaagtgtcagga 1461
 DB 1405 TCTGCTCACTGTCTCAATTTATCTATGACATTAATTAAGACAAATAATCCAGTGGCAGGG 1464
 QY 1462 aaggaacaccaggtgctcctgaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1521
 DB 1465 AAGAAACATCAGAGTGGCTGT 1524
 QY 1522 gt 1581
 DB 1525 TTGGTATGGGTTCACAGCTGCTCTTCAAGTGTGATGATGATGATGATGATGATGATGAT 1584
 QY 1582 ctcaacaagaagtgagacttcaagatgccaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1641
 DB 1585 CTCACACAGAGTGAAGCTTCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644
 QY 1642 ggaagcagagcgt 1701
 DB 1645 GGGAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1704
 QY 1702 tgggt 1761
 DB 1705 TTGGTCTATTTAGACACTTAACAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1764

QY 1762 ttgacctggcaccctagacacacccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1821
 DB 1765 TTGCTGGGCTGCTTCAACAGCTGTTGTGAGGTGACAGGGGGGTGAGGCTGCTCTC 1824
 QY 1822 atgt 1881
 DB 1825 ATGCTGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1884
 QY 1882 aggcctgt 1941
 DB 1885 GTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
 QY 1942 tgcctgacagctgt 2001
 DB 1945 TGTCTGACATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2004
 QY 2002 acatttaccacagcctgt 2061
 DB 2005 ACATTTACACACTTGGGCGCAAAACATGCGCGGAAATATGCTCATTTGTCACTCTC 2064
 QY 2062 gggccagctgt 2121
 DB 2065 ACGGTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2124
 QY 2122 gaataccagcgt 2181
 DB 2125 GATGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2184
 QY 2182 ttcatatgt 2241
 DB 2185 TTGCTGT 2244
 QY 2242 ctgggtaagagctgt 2301
 DB 2245 CTGGGTAAAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2304
 QY 2302 ttcaactgt 2361
 DB 2305 CTCACATTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2364
 QY 2362 ctgacctgt 2421
 DB 2365 CTACCGCGGCTGATGT 2424
 QY 2422 ttctgtcctaaagt 2481
 DB 2425 TTGCTCTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2484
 QY 2482 cagacctcattcagagctacagagagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2523
 DB 2485 CAGGCTTCATCCAGAGCTACACAGAGGCGCTGCGGCACTAAC 2526

RESULT 6
 AF337040 2892 bp mRNA linear ROD 25-APR-2001
 LOCUS AF337040 Mus musculus candidate taste receptor T1R1 (T1R1) mRNA, complete cds.
 DEFINITION AF337040
 ACCESSION AF337040
 VERSION AF337040.1 GI:13785656
 KEYWORDS
 ORGANISM house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 AUTHORS Montmayeur, J.-P., Liberles, S.D., Matsunami, H. and Buck, L.B.
 TITLE A candidate taste receptor gene near a sweet taste locus
 JOURNAL Nat. Neurosci. 4 (5), 492-498 (2001)
 PUBMED 11319557
 REFERENCE 2 (bases 1 to 2892)
 MONTMAYEUR, J.-P., LIBERLES, S.D., MATSUNAMI, H. and BUCK, L.B.
 Direct Submission

JOURNAL Submitted (18-JAN-2001) Neurobiology, Harvard Medical School, 220
Longwood Avenue, Boston, MA 02115, USA
FEATURES
source location/Qualifiers
1. 2892
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="4"
1. 2892
/gene="Trlr1"
183. 2711
/gene="Trlr1"
/note="G protein coupled receptor"
/codon_start=1
/product="candidate taste receptor Trlr1"
/protein_id="AAK39437.1"
/db_xref="GI:13785657"
/translation="MLEMAHLLLSIQLAAYAWAFSCQTESSPEFSLPDELACL
FSLHADCLQVRRPLVYSCDRSDSFNGHGLFOAMFTEEINNSALLPNTLIGE
LYDVCSSESNVATLRYLAQOGTHLEMODLNHSSKVVALLGPDTHAVTAA
LSPRLPLVSEASSVILSGKRKFPFRTIPSDKOYEIVRLLOSGLWMSVLSY
GDYGLQVQALELAPRGICVAFKDVYPLSAQADPRMRMLRLARATYVYVS
NRHAGYFRSVVLANLTKVWLASEDMALSTYTPG10G1GVLGVA10QRQVPG
LKEFESYVQAVMGAPRTCEGSMGCTNOLCRCHAFETNMDELAFSMAVNYE
AVYAVAGHLHQLIGCTSGTCARGPVYPMQLQOIKYVNFLLHKKVAFDQDPLGY
DIAMPNGEEMTFEVIKASISPVHLIDIKTKIOWHGNNOYVPSVCTDCEGHR
LVMSHHCECEMPCAGTFLNSLHLOPCGTEEMAPEGSSACEFRTVEFLGHRP
ISVILAAANTLLLLIGTGLFAMRLHRTVYVSAGRLCFMLGSLVASSCSISXF
GKTPVACLRLQPLSLGFAIFLSCILTSFOLVILFKESTKVPTEHTWAQNHGAI
FVIVSVHLFLCLTWMAMTPTREYRPHVILIECTEVSVGLVAFANHLIS
ISTFVCSYLKELPENYNAKCVTFSLHFEVSWIAFEFTSSITQGSYLPVAVNLAL
ATLSGSGSFLKCYVILCRPELNTEHFOASIDYTRRCGTT"

BASE COUNT 585 a 841 c 794 g 672 t
ORIGIN

Query Match 66.9%; Score 1691; DB 10; Length 2892;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 1994; Conservative 3; Mismatches 508; Indels 0; Gaps 0;

QY 22 ctgttcgagcttgacgttcaattcctctgctgctgagccttgcctgcataagacgag 81
DB 207 CTGCTCAGCCGTCAGCTGGCGGTGCTTACTGCTGGCGTTTCAGCTGCCAAGACAGAA 266
QY 82 tctctctactacacccctcccgagattactcctcggagagcctggttccctcctcat 141
DB 267 TCTCTCCAGGTTTCAGCTCCCTCTGGGACTTCTCTGGGAGCGCTTCTCCCTCAT 326
QY 142 tctgagctctgacgttgagacacagaccgagtgaccctggttgacaggtcttgagc 201
DB 327 GCTGACTGTCTGCAGGTGAGACACAGACTCTGCTGACAAAGTGTGACAGGTCTGACAGC 386
QY 202 ttcaatgacatgctacacaccccttcccaagctatcgagccttgaggttgagagataaac 261
DB 387 TTCAAGGCGCATGCTGCTTCACTCTTCCAAAGCCATGCGGTTTCACCGTTGAGAGAGATAAC 446
QY 262 aaccacacgagcctgctgcccacaatacaccctgggtgtacacagcgtatgagtgtgt 321
DB 447 AACCTCAGAGCTCTGCTCCAAACATCACTCGGAGTATGAACTGTATGAGCTGTGCTCA 506
QY 322 gactctccaaatgtgtatgacagctgagagtgctctcctcctgcagaggaacacacata 381
DB 507 GAGCTTCCAAATGCTATGACCACTGAGGCTGCTCGCCGACCAAGGACAGGACACCTTA 566
QY 382 gagctcgaaggaacacctctccactatccctacggtgtgacagtgagtgagtgagc 441
DB 567 GAGATGAGAGAGATTTTCGAACACACTCTCCAAAGGTGAGGACATCATTTGGGCGCTAT 626
QY 442 agaccacacgctgtgacacacagccgctctgtagagccttccctgggtgacatgatt 501
DB 627 AACACTGACACGCTGTCTCAGCTGCTGCTGACGCCCTTTTGTGATGCCCTTGCTC 686
QY 502 agcatatgcgcagacagagacgctcagcgttgagcgagatatacccttctcctgagc 561

DB 687 AGCTATGAGGCGAGCAGGCTGATCCTCAGTGGAGGCCAAGTTCCCTCTTCTGGCC 746
QY 562 accatccccaatgacaaataccaggttgagacacatgctgctgctgctgcagaagtccg 621
DB 747 ACCATCCCCACGATTAATGATCAGGTGAGAACTATGAGCGGCTGCTCAGAGCTTCGGC 806
QY 622 tggaccctgacatctctgcttggtgagcagctgagacatactggagcagctgagtgagca 681
DB 807 TGGGCTTGATCTGCTGCTGTTGGCAGCTATGCTGACATCGGCGAGCTGGGCTACAGCG 866
QY 682 ctggagaaccagcagctgctgacagggagatctgcatctgcttccaaaggaatcatgacctc 741
DB 867 CTGAGAGAGCTGCGCACACTCCACGGGGCATCTGCGTGGCTTCAAGAGAGTGGTGGCTTC 926
QY 742 tctgcccagctggtgagatgagagatgcatgctcctcatgctgcccacctgcccagccg 801
DB 927 TCCGCGCCAGGCGGTGACCCCAAGGATGACGCGCATGAGCTGCGCTCGAGCCAGG 986
QY 802 gccacacgctgctgctgcttctccagccgagctgagcagaggtgttcttcagatccgtg 861
DB 987 ACCACCGTGTGCTGCTGCTTCTTACCGGCACTGCTGGAAGTCTTTCAGCTCTGTG 1046
QY 862 gtgtgacaaacctgactgagcaaggtgtggtgcgctcagaagcctgagccctccag 921
DB 1047 GTGCTGGCCAACTGACTGCGCAAGTGTGATCCGCTCCGAAGACTGGGCCATCTCCAGC 1106
QY 922 cacatcaatggtgtgcccggatccagccgcatgtgagatgctgctgctgctgctgcatccag 981
DB 1107 TACATTCACCAATTTGCCCCGGGATCCAGGGCATTTGGAGGTTGGGGGTGGCCATTCAG 1166
QY 982 aagagagctgctcctgagcctgaagcgtttgaagaagcctatgcccggagcagacaagag 1041
DB 1167 CAGAGACAAGTCCCTGGCTTAAGGAGTTTGAAGATCTTATGTCAGAGCAGTATGGGT 1226
QY 1042 gccctttagccttgccacaagagctcctctgtgacagcaatcagctctgcagagaatgc 1101
DB 1227 GCTCCCAAGACTTGGCCAGAGGGGCTCTGTTGGGCACTAACACAGCTGCGAGGAGTGT 1286
QY 1102 caagcttcatatgacacacagatgccagctcaagccttccatgagatctgctgctac 1161
DB 1287 CACGCTTTCAGACATGAGAACATGAGCCGAGCTTGGAGCTTCTCATGAGCGCTGCCTAC 1346
QY 1162 aagcatatcgagctgtgtatgacggtgagccatgagccctcaacagcctctgagcctgagc 1221
DB 1347 AATGTTATGAGGCTGTATGCTGTGCTGACCAAGGCTCCACAGCTCTGGAGTATGCC 1406
QY 1222 tctgagcctgttccagagggcagcttaccctctgagccttctgagcagatccacaag 1281
DB 1407 TCTGGAGACTGTGCGACAGAGGCCAGTCTACCCCTGCGAGCTTCTTCACAGCATCTACAG 1466
QY 1282 gtgaattccttctacacaaggaacatgtagcgtttaaatagaacaagaagatccctcagt 1341
DB 1467 GTGAATTTCTTCAATAAAGAAAGACTGTACATTCGATGACAAAGGGGAGCCCTCTAGGT 1526
QY 1342 agctataacataattgctgtagactgtagaattgagcccaagtgagcttcaagtgctcgtg 1401
DB 1527 TATTATGACATCATGCTCTGGAGCTGAGTGAACCTGATGAGCTTGAAGTCAATTTGT 1586
QY 1402 tccctcacaatgcttccagctcagctcagctaaataatgagacaaatccatgagcagga 1461
DB 1587 TCTGCTCTACATGCTCTCCAGTTCACTACATTAATAAGAAAAATTCATGCGACGG 1646
QY 1462 aaggaacaacagctgctcagatctgctgctgctcagcagctgcttctgaaggacacagga 1521
DB 1647 AAGAAACATATCAGTGGCTGTCTGATGATGACAGGGAGCTGTGCAAGGCGACACAGG 1706
QY 1522 gtgtttagcgggttccatcaatctgctttagagtggtgagcccttgaggagccttc 1581
DB 1707 TTGGCTATGAGGTTCCACACACTGCTGCTTGAAGTGCATGCCCTGTGAAGCTGGGACATTT 1766
QY 1582 ctcaacaagaatgacatctacagatgacagcctgtgtggaagaagaatgagcagctgag 1641

OY	382	gagctcaagaagagacctcttcacatcttcacctgaagtgctgcagtgatctggagctgcac	441
Db	385	GAGATGCGAGAGAGATCTTCCCAACCACTCCTCCCAAGTGGTGGCACTCATGGGCGCTGAT	444
OY	442	agacacaaaccgtctgcaccacaagccgcctctgtagacccttcctgtgtcccaatgt	501
Db	445	AACCTGACCAACGGTGTCCACCACTGCTGCCCTCTGTAGCGCTTTCTATATGCGCCCTGGTC	504
OY	502	agctatcgagccagcaacgcagacgcctcaacgcgtgaagcgcgaatctccctcttcctgcgc	561
Db	505	AGCATATAGCGGAGACACACCGATGATCTCTCATGTGGGAACGCAATTTCCGCTCTTGGCGC	564
OY	562	accatccccaatgacaagtaacagatgtagaacatgtgtgtctgtctgcagaaagltgcgg	621
Db	565	ACCATCCCCAGCGGATAGTATCCAGAGTGGAAGTATAGTGCGGCTGTCGAGAGCTTTCGGC	624
OY	622	tgagaccgtgatatctctctgtgtgtgcagtgaaagataatgggcagctaggggtgcagga	681
Db	625	TGGGTCTGTGATCTCGCTCGTGTGGCACTATGTGATAGGGGCACTGGCGCTTACAGCGC	684
OY	682	ctgtagaaccaagcagccacttggtaaggagatctgatgtgtcttcaagaataatgaccttc	741
Db	685	CTGAGAGAGCTGGCGCACTCCACGGGGCATCTCGTGGCTTTAABAGAGTGGTGGCTTTC	744
OY	742	tctgcaccaagtgtaggcgataggaagatgacgtgctcatatgcacactgtgccagagccgg	801
Db	745	TCCGCGACGGGGGGGTGACCCCAAGGATGACAGCGATGATCTGCTGGCTGAGACGAG	804
OY	802	gcacacgtctgtgtgttcttctccagccggacagttgtgccaagtglttcttgagtctgt	861
Db	805	ACCACCGTGGTCGGGTCTTCTTAACCGGCACTGGCGAGTGTCTTCAGGCTCTGTG	864
OY	862	gtgtctgaccaacactgactgtgcaaggtgtgtgtgcctcaagaagccttgagccctccag	921
Db	865	GTCGTGGCCCAACTCGACTGGCAAGTGTGATGCTCCGAAGACTGGGCCATCTCCACG	924
OY	922	cacatactcagtggtgtcccggaatccagcgcgcatgtggaatgtgtgtctgtggcgtgcac	981
Db	925	TACATACCAATGTCGCCGGGATTCAGGGGATTTGGGACGCTGTGGGGGTGGCATTCAG	984
OY	982	aagaggtcgtctccctgtgcctcgaagcgcgtttgaagaagcctatgtcccggtcagaagaag	1041
Db	985	CAGAGACAAAGTCCCTGGCCCTGAAGGAGTTTGAAGACTCTATGTCAGGCACTGATGTGGT	1044
OY	1042	gccctcagggccttgccaagaaggtcctctgtgtgacagacaaatcagctctgcagagaatgc	1101
Db	1045	GCTGCCAGAACTTCCCAAGAGGGGCTCTGTGGCGCACTAACCACTGTGCAAGGAGTGT	1104
OY	1102	caagcttctcatgiraacacagatgtgccaaagctcaaaagccttctcatatgattctgtcac	1161
Db	1105	CAGCTTTCAAGAAATGAGAACATGCCCCGAGCTTGGAGCTTCTTCATGAGCGCTCTCAC	1164
OY	1162	aaagcataccgggtctgtgtatgtgcgtgtggccatgtgcctccacagctcctgtgctgtgc	1221
Db	1165	AATGTGATGAGGCTGTGTGTGTGTGGCCACAGGCGCTTCACACACTCTGTGGAGTATCC	1224
OY	1222	tctgtgagctgttctcaagggtgcaggtctcaaccctgtgcagcttttgagcaaatccaacag	1281
Db	1225	TCTGTGGAACCTGTGCGCAGAGGCCCAAGTCTAACCCCTGGCACTTCTTCAGCAATCTACAG	1284
OY	1282	gtgcatttctctctacacaagaacactgtgtgcgttaatgacaacagaagatccctcact	1341
Db	1285	GTCGATTTCCCTTACATTATAGAGACTGTAGACTTTCGATGAGCAAGGGGACCTTATAGT	1344
OY	1342	agctataacataatgtgcctgtgcgttggaatgtgaacccaagtgtgaaccttaacggtctcgt	1401
Db	1345	TATATATGACATCAATCGCTGGACTGGAATGAGACCTGAATGAGACCTTTGAGGTCTATTGGT	1404
OY	1402	tcctccacaatgtgtctcagctcagctcaaaacataaagtgaagacaacaaatccaatgtggaag	1461
Db	1405	TCTGCTCATCAGTGTCCAGTTTCACTTAGACATATTAATTAACAAAAATTCAGTGGCAACGG	1464
OY	1462	aaggacaaccaagtggtccctaagctgtgtgttctcagagcactgtcttgaaggagacaacgga	1521

Db	1465	AAGAACATCGAGGTCCTGTGTAGTGTACAGGAGACTGTCTGAAAGCGACCAACAGG	1524
QY	1522	gtgttaaggggtttccaatacgtcgtttgaatgtgtgtccctgttgggcttggagaccctc	1581
Db	1525	TTTGTCATGAGGTATCCCAACCACTGCTGTGATGATGCAATGCCCTGTGAAGCTTGGACATT	1584
QY	1582	ctcaacaagaagtgaacctctacaagatgcgaacctgtgtgggaaagaagatgvggcaacctgag	1641
Db	1585	CTCAACACAGAGTGTAGCTTTCAACACTGCGACACTTGTGTGAACAGAAAGAAATGGGCCCTTAG	1644
QY	1642	ggaagccagagaccbctctcccgagcactgtgtgttttggcttcttggtgtgagcaacctc	1701
Db	1645	GGGAGCTCAGGCTCTCTCTCAACGACCGTGTGAAGTTTGGGGTGGGATTAACCCATCTCT	1704
QY	1702	tgagggtgcctcgtgagagtaacaagacgtcgtcgtcgtcgtcgtcgttgaggagacgtcgtg	1761
Db	1705	TTTGCTGCTATTAGACAGCTTAACAGCGTATTTCTCTCTGCTGATTTGGACTCTTGCCCTG	1764
QY	1762	tttgcctgtgcacccctacacacccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1821
Db	1765	TTTGGCTGTGGCTTTTTCACACGCGCTGTGTGTGAGGTGTAGGTGGGGTGTAGGCTGTGCTTC	1824
QY	1822	atgctgtggctctccctgtgcgaagaagtgtgtgcagcctctatgtgctcttcttggggaaaccaca	1881
Db	1825	ATGCTGGGTTCCTTGTGTGAGTGTGGAGTTGGCAGCGCTTACAGCTTCTTCGGGAAGCCACAG	1884
QY	1882	agggctgtgtcgttgcctgaagcagggccctcttggccctgtgtgtgtgtgtgtgtgtgtgt	1941
Db	1885	GTGCGCGGTCTGTCTGTGCGCTGACGCCCTTTTTCCTGTGGGTTTGTGCCATTTTCTCTTCC	1944
QY	1942	tgccgtgaacagttgcctatctaccataatcatcatcttaagtlttccacaagttacct	2001
Db	1945	TGTGTGAACAATCCGCTCTTCCCAACTGTGTATCATCATCTTCAAGTTTCTACCAAGSTACCC	2004
QY	2002	acattctaaccaagcgtgtgtgtccaaacaagtgctgtgcgtgttgtgtgtgtgtgtgtgtgtgt	2061
Db	2005	ACATTCTACCAACTGTGGGCCCAAAACAAGTGTGGGAATATCTCATTTCTGACGTCC	2064
QY	2062	ggcgcccaagcgtctatctcttcaacttggctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2121
Db	2065	ACGGTCCATTGTTGCTCTGTCTGTACGTGGCTTGTGAATGTGACCCGACGCCACGACACAG	2124
QY	2122	gaataccaagcgtctccccaatctgt	2181
Db	2125	GAGTACAGAGCCCTTCCCCCACTGTGTATTTCTGTGAATGTGACAGAGGTCAACTGTGTGGC	2184
QY	2182	ttcatactgtgctctcctctacaatgtgctctctctcatgaatgtgcttctgtcgtgagctac	2241
Db	2185	TTCTCGGTGGCTTTTGTGCAACAACAATCTCTCTTCATAGCACTTTGTGTGCAAGCTAC	2244
QY	2242	cttggttaagagatctgtgcagaagaactacaagaagggccaatgtgttcaacttcaagcctgtctc	2301
Db	2245	CTGGGTAAAGAACTGTGGGAGAACTAATAAGAAAGCAATATGTGTACCTTCAAGCCTGTCTC	2304
QY	2302	ttcaactctgtctctcctcctcctcctctctcaacaagggcaagcgtcttaegaagggcaagtac	2361
Db	2305	CTTCACATTCGTATCTGTGATGCGCTTTCTTCACATGTTCAGCAATTATACGAGGGCAGCTTAC	2364
QY	2362	ctgctctgtggccaacaagatgtgcttgggtcttgaagcctctgaagaagcgggtctgtgtgtgtat	2421
Db	2365	CTACCCGGGGTCAATGTGTGTGGCAGGGGCTGTGCCACTCTGTAGTGTGGCGGCTTCAAGCGGCTAT	2424
QY	2422	ttctgtcccttaagtgtctagatctctctgtccgccaagagactacaagcagaagaagcctc	2481
Db	2425	TTTCCTCCTTAATCTTACGTGATTTCTCTCCGCTTCAGAACTCAACAACACAGACACTTT	2484
QY	2482	cagggctccatctcaagactacaagaagcgtctgcggtcccaacctgtga	2526
Db	2485	CAGGCTTCATTCAGGACTACAGAGCGCTGCGGCACTACACGTA	2529

RESULT 8

AF301161 2526 bp mRNA linear ROD 12-FEB-2001
 LOCUS AF301161
 DEFINITION Mus musculus strain C57BL/6ByJ putative sweet taste receptor T1R1
 (Gpr70) mRNA, partial cds.
 ACCESSION AF301161
 VERSION AF301161.1 GI:12745517
 KEYWORDS house mouse.
 SOURCE Mus musculus.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2526)
 Li, X., Inoue, M., Reed, D. R., Hugue, T., Puchalski, R. B., Tordoff, M. G., Nimomiya, Y., Beauchamp, G. K. and Bachmanov, A. A.
 High-resolution genetic mapping of the saccharin preference locus (Sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to mouse distal Chromosome 4
 Mamm. Genome 12 (1), 13-16 (2001)
 JOURNAL 21030739
 MEDLINE 2 (bases 1 to 2526)
 REFERENCE Li, X., Reed, D. R., Hugue, T., Puchalski, R. B., Tordoff, M. G., Beauchamp, G. K. and Bachmanov, A. A.
 AUTHORS Direct Submission
 TITLE Submitted (29-AUG-2000) Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104-3308, USA
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..2526
 /organism="Mus musculus"
 /strain="C57BL/6ByJ"
 /db_xref="taxon:10090"
 /chromosome="4"
 /map="between D4Mit42 and D4Mit254"
 1..>2526
 /gene="Gpr70"
 1..>2526
 /codon_start=1
 /product="putative sweet taste receptor T1R1"
 /protein_id="AAK07091.1"
 /db_xref="GI:12745518"
 /translation="MFMFAHLLSLQLAAYCAAFSCORTSSPQSLPDELALG
 FSLHDCLOVRRLVLTSCDRSDSRNGHYHLFQAMRFLVEEINNTALLPNTTIGYE
 LTVDSSESNVATLRLVLAQGTGLHEMDRDLNNSKVALIGDNDHVAVTTAALL
 SPFLMPLVSEASVTLISGKRKPSFLRTIPSDKYQEVIVLILQSGFMDVAVTLLVGSY
 GDYGLGVQALELALPRGICVAFKRVVLSAQAGPRQRMRLRLARARTVYVVS
 NRHLGVFRSVLVANLNGKVIASPDMAISINIPVSGIOGIVLVAIOORVPG
 LKEFESYVOAYMGAAPICPREGSCGTNOLRECAFTMNMPELGAFSMAAYNVE
 AYVAHGLHQLGCTSGTCAGPVPHPMLOLQITKYVNLHLKRTVADPKDGLGYT
 DIIAMDWNPENTFEVYIGSASLSPVHLIDINKRTIQHGNVNPVPSVCTRDLEGHHR
 LVMGSHHCCFECMPCAGFLNLTSELTDCPCGTERMAPGSSAGFSRVEFLGHEEP
 ISLVLLAANTLILLILIGTAGLFLAMRLHTPVVRSAGRLCFMLGSLVSGSLVSEF
 GKPTVACLLRQPLFLSGLFALISCLTISFQDLVITFKSTKVPFPHYHMAONHNGI
 FYIVSIVHLFELCTLWAMPMPRPRTREYORPHVILLECTEYNSVGFYAPAHNLLS
 ISTPVSYIAGKELPENYNEAKCVTSLLHFYSWLAFTFMSSTIYQSYLPAVNVLAGL
 ATLSGFSFYFLICREPLNMTHEFOASIDYTRRCGTT"
 BASE COUNT 485 a 755 c 679 g 607 t
 ORIGIN
 Query Match 66.4%; Score 1678.4; DB 10; Length 2526;
 Best Local Similarity 79.3%; Pred. No. 0;
 Matches 1985; Conservative 3; Mismatches 514; Indels 0; Gaps 0;

Db 145 GCTAGCTGTCTGAGGTGACACACAGACCTCTGTGACAAAGTTTGACAGAGCTTGACAGC 204
 Qy 202 ttaatgagatggtctaccaccctccctccaggtatgcggttgagagataaac 261
 Db 205 TTCAAGGGCCATGGCTATACCTCTTCCAAAGCATGCGGTTACCGCTTGAGAGATAAAC 264
 Qy 262 aactccagcgctctgtcccaacatacccttgggttaccagctgatgatgtgtc 321
 Db 265 AACCTCAGACCTCTGCTCCACATCACCTCGGGTATGAACTGATGACGCTGCTCA 324
 Qy 322 gacttcgcaatgtgatgtccagctgagatgtctctccctccagggcaaacccata 381
 Db 325 GAGCTTTCATGTCTATGTCACCTGTGAGGCTCTGCGCAGAAAGGACAGGCACTTA 384
 Qy 382 gaactccaaagaaaccttccactatccctccacagtgctgcagtgatgtggcctgac 441
 Db 385 GATATGACAGAGATATCTCCGCAACCACTCTCCAAAGTGGTGACATCATTTGGCCTGAT 444
 Qy 442 agccaacacgtgtcccaacacagccgcctgtcgaagcccttctgtgtccatgatt 501
 Db 445 AACACTGACACGCTGTCTACACCTGCGCTGTGAGCCCTTTCTGATGCCCCGTGTC 504
 Qy 502 agctatgcgccagcgagcgagacgtctcagctgaaagcgagatccctcttccgcgc 561
 Db 505 AGCTATGAGCGCAGACCGCTGATCTCAGTGGGAAGCGCAAGTTCCGCTCTTCTTGCGC 564
 Qy 562 accatcccaatgacagatccagatgagacatgtgtcgtctgtcagaaatlcggg 621
 Db 565 ACCATCCCGACGATATAGTATCCAGGTGGAAGTATAGTGGGCTGTGCGACACTTCGCG 624
 Qy 622 tggactgatatctctgtgtgtgagcagatgacgactatgagcagctaaaggttgcaag 681
 Db 625 TGGGTCTGATCTCGCTCTGCTGTCGACCTATGAGTACAGGCGAGCTGGGCGCTACAGCG 684
 Qy 682 ctggaagaccagcgccctgtgacggagatctgattgtctttaaagacataagccctc 741
 Db 685 CTGGAAGAGCTGGCTCACCTCACGGGATCTGTGCTCTTCAAGAACGGTGCTCTCTC 744
 Qy 742 tctgcccaggttgagcatgagagatgacgtctcatctgcacacctggccagcgcg 801
 Db 745 TCGGCCACAGCGGGTACCCAAAGATGCAAGCGATATGCTGCTGTGCTGACGACAG 804
 Qy 802 gccacgctgtgtgttcttccagcggcagctgtgccaaggtgttctgaatccgtg 861
 Db 805 ACCACGCTGTGCTGTCTCTCTTCAACCGGACCTGAGTGAATGTTCTTCAAGTCTGTG 864
 Qy 862 gtgtgacaaactgtctgcaagatgtgtgtgtccttcgaagccctgggccccttcag 921
 Db 865 GTGCTGGCCAACTGACTGGCAAAAGTGTGATGCTCTCCGAAGACTGGGCCATCTCACG 924
 Qy 922 caatcaactgggtgtcccgagatccagcgcatgtgagatgtgtgtggtgtggtccatccag 981
 Db 925 AACATCCCAATGTGTCCGGATCCAGGCGATTGGGACGGTGTGGGGGTGGCATTCAG 984
 Qy 982 aagagggctgtccctgtgctgaagcgcttgaagaaacctatgtcccgagcaagaagag 1041
 Db 985 CAGAGACAAGTCCCTGCGCTGGAAGGATTGAAGACTCTATGTCCAGGACAGATGAGGT 1044
 Qy 1042 gcccttaggccttgcacaaaggtctctgtgtgagcagcaatcagcttcgagagaatgc 1101
 Db 1045 GCTCCCAAGAACTTCCACAGAGGGTCTGTGTGGCGACTAACCACTGTGCGAGGGTAT 1104
 Qy 1102 caagcttatgatacacacagatgcccagaagcttcaagcttccatgagcttctgtctac 1161
 Db 1105 CACGCTTTACGACATGGAACATGCTCCGAGCTTGGAGCTTCTCCATGAGCGCTGCTTAC 1164
 Qy 1162 aacgcaacacgggtgtgtatgctgtgcacatgagctccacacagcttccgtgtgtgc 1221
 Db 1165 AATGTGTATAGAGCTGTGATGCTGTGGCCACAGGCTCTCACAGAGTCTCTGGAGATAC 1224
 Qy 1222 tctgagactgtgtccagggcgagctctacccctgtgcagcttcttgagcagatcatcaag 1281
 Db 1225 TCTGGGACCTGTGCCAGAGGCCAGTCCACCTTGCGACCTTCTTCAGCAGATCTTAAG 1284

QY 1282 gtgattctctctacaaagacagtigggcfttaatacaacagatccctcaagt 1341
 DB 1285 GTGATTTCTTCTACTAAGAAAGACTGTAGCTCGATGCAAGGGGACCCCTAGT 1344
 QY 1342 agctatacaataatcttgctggactggaatggaaccagaatggaactcagctcgt 1401
 DB 1345 TATTATACATCATCGCTGGAGCTGGAATGGACCTGAATGGACCTTTGAGGTATGCT 1404
 QY 1402 tctctcaaatggtctccagttcagctaaataaataagaaacaaatccagttgacgga 1461
 DB 1405 TCTGCTCCTACATGCTCCAGTTCACTAGACATAATAGCAAAAATCCAGTGCGACGGG 1464
 QY 1462 aaagacaagagctgctgaagctgtgtgttccagagcagctgtggaaggaccacgga 1521
 DB 1465 AAGAACATACAGGTGCTGTGTGCTGAGTGTACAGGAGACTGTCTCGAAGGACACACAGG 1524
 QY 1522 gtgttcaaggttctccatcactgctgcttgaagtgtgtgctgctggaagctcagcttc 1581
 DB 1525 TTGGTCAATGGGTTGCCACCACTGCTGCTGAGTGCATGCCCTGTGAAGCTGGACATTT 1584
 QY 1582 ctcaacaagagtgaacctctacagatccagcctgtgtggaagaagaagtgggaacctgag 1641
 DB 1585 CTCAACACGAGTGAAGCTTCAACACCTGCGACGCTGTGGAACAGAAAGATGGCCCTGAG 1644
 QY 1642 ggaagacaagagctgctccgagcagctgtgtgttctgctgtgtgagcagacactct 1701
 DB 1645 GGGAGCTACACCTGCTTCTTACACGACCGTGGAGTCTTGGGGTGGCATGAACCCATCTCT 1704
 QY 1702 tggagtgtctgctgcaagtaaacacgctgctgctgctgctgctgctggaagctgctgct 1761
 DB 1705 TTGGTGTATTTAGAGCTTAACAGCTATGCTGCTGCTGCTGAGATTGGAGCTGGGCTG 1764
 QY 1762 ttgctgtggaacctagaacacctgtgtgtgaagtcaagaggggcgcgtgtcttctt 1821
 DB 1765 TTTGGCTGGGCTCTTCAACAGCTGTTGTGAAGTCACTGGGGTGAAGCTGTGCTTCTC 1824
 QY 1822 atgtgtgctccctgcaagcagagtagtggcagcctctatgctcttctggaagacacaa 1881
 DB 1825 ATGCTGGGTTCTTGTGAGCTGGAGTGGACGCTTACAGCTTCTTGGGAAACCCACG 1884
 QY 1882 agagctgtgtgtctgtcaagcagagccctcttgccttggtttaccactcttctgccc 1941
 DB 1885 GTGCCCGGAGCTGTGTGCGACAGCCCTCTTCTCTCGGGTTTGGCATTTTCTCTCC 1944
 QY 1942 tgcctgacagttgctcattccaaataataataatcttcaagtttccaaagtaacct 2001
 DB 1945 TGTCTGACAAATCCGCTCTTCCAACTGTATCATCTTCAAGTTTCTTACCAAGGTACC 2004
 QY 2002 acattcaacaagcgtgtgtccaaacacagctgtgctgtgtgtgtgatcagctca 2061
 DB 2005 ACATTTCAACCACTTGTGGGCCAAAACCAATGTCGCGAATATTCGTCATTGTGACGCTC 2064
 QY 2062 gcggccagctgtcttactgtcttcaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2121
 DB 2065 ACGTTCATTTGTTCTCTCTCTCTCACTGCTGCTGCAATGTGACCCACGCGCCACACG 2124
 QY 2122 gaatacaagagcttcccccatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2181
 DB 2125 GAGTACACAGGCTTCCCTCCATCTGTGATCTTGTAGTGCACAGAGGTCAACTCTGTGGGC 2184
 QY 2182 ttcatactgctcctctccataatgagcctcctccataagttcctgctcctgctcagctac 2241
 DB 2185 TTCTGTGTGGCTTTCGACACACATCTCTCTCTCATCAGACCTTTGTGTGACGCTAC 2244
 QY 2242 ctgggttaagagctgtgcagagaactacaagagggccaatgtgtcacttcaagcctgtc 2301
 DB 2245 CTGGGTAAAGAACTGCGGAGAACTATACGAAGCCAAATGTGACCTTACGCTGCTC 2304
 QY 2302 ttaactgtgtctgtgctgctcttcttcaacagcgcagcgtctcagagaggaagtac 2361
 DB 2305 CTGCACTTGTATCTGTGATGCTTCTTCTTCACTGTGACGATTTACAGGGGAGCTAC 2364

QY 2362 ctgctcgcggccaacatgatgtcgtgtgagcagcctgaagcagcgtctcgtgtgtat 2421
 DB 2365 CTACCCGGGTCAATGTGCTGTGCGAGGCTGGCCACTGAGTGGCGGCTTACGGGCTAT 2424
 QY 2422 ttctcctaagtctcagttatcctctgcccacacacacacacacacacacacacacacac 2481
 DB 2425 TTCTCCCTTAATGCTACGTGATCTCTGCGCTCCAGACACTCAACACAGAACACTTT 2484
 QY 2482 caggctccatcagactacagagcgcgtgtggtccacc 2523
 DB 2485 CAGGCTTCATCCAGACTACACAGGCGCTGCGCACTTACC 2526
 RESULT 9
 AF127389 2520 bp mRNA linear ROD 04-MAR-1999
 LOCUS AF127389
 DEFINITION Rattus norvegicus putative taste receptor TR1 mRNA, partial cds.
 ACCESSION AF127389
 VERSION AF127389.1 GI:4337085
 KEYWORDS
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 2520)
 Hoon,M.A., Adler,E., Lindemeier,J., Battey,J.F., Ryba,N.J. and
 Zuker,C.S.
 TITLE Putative mammalian taste receptors: a class of taste-specific GPCRs
 with distinct topographic selectivity
 JOURNAL Cell 96 (4), 541-551 (1999)
 MEDLINE 99159821
 REFERENCE 2 (bases 1 to 2520)
 Hoon,M.A., Adler,E., Lindemeier,J., Battey,J.F., Ryba,N.J.P. and
 Zuker,C.S.
 TITLE Direct Submission
 JOURNAL Submitted (10-FEB-1999) Taste and Smell Unit, NIDCR, 10 Center
 Drive MSC 1188, Bethesda, MD 20892-1188, USA
 FEATURES
 source location/Qualifiers
 1..2520
 /organism="Rattus norvegicus"
 /strain="Wistar"
 /db_xref="taxon:10116"
 /cell_type="Circumvallate papilla"
 1..>2520
 /note="G protein-coupled receptor"
 /codon_start=1
 /product="putative taste receptor TR1"
 /protein_id="ADI8069.1"
 /translation="MLFMAHLLLSLQVLCMAFSCORTSSPGLPDPFLAGLES
 LHDDCLQVRRLPVTSQDRPSNGHYNLFQMRKRVATETINSSALIPVITGLYEL
 DVCSANVATATLVLAQGPRIHLEIQRDNRHNSKYVAITGPDNDHTATLALQ
 FLMPVLSYEASVYLSAKRPFSLRTPSDRQVEVMVQLQSGVWVLSLGSYGD
 YGQLEQVLAELAVPRGICVAFEDIVFSAVDPRQSMQHLAQARTVVVFSNR
 HLAIVEFSAVLANLNGKVVVAEDMAISYITSVGIQIGVLAQVQROVPGKL
 EFEEYSYRAVTAAPSACRSGSSTNQLQCEHPTTRNRPITGASMSAAYEAV
 YVAHGHQLGCTSEISGRPYRPMQLQIYKVNFLHENTYAPDNDPTLYXDI
 IAMDNGPEMTFETIGSASLSPVHLIDNKRITOMHGNNQVPVSVCTTDLACHHRV
 VGSNHCQFECVPCPEAGFTLMSLHLICQPGTEWAPKESTTCEPRTEVLAMHEPIS
 LVLIAANTULLLVAGTAGLFAWHFHPVVRSAAGRLCFMLSLVAGSCEFSEFGE
 PVPACILRQPLSLGFAFLPSCLTISFQDLVLIKFSTKVPFVPTMAQNHGALFV
 IVASTYHLILCLMLVMMTPRPRREVORPRLVLIETEVNSVGFLLAPFHNILSIS
 TPVCSYIGKLELPENYNAKVTYSLLNTPSNIAFPMASIIYQGSYILPAVNVLAGLTT
 LSGFSYGLFPLKCVILICRPELNWTEHFOASIDYTRKCGIT"
 BASE COUNT 479 a 749 c 676 g 616 t
 ORIGIN
 Query Match 66.28; Score 1671.4; DB 10; Length 2520;
 Best Local Similarity 78.96; Pred. No. 0;
 Matches 1987; Conservative 3; Mismatches 529; Indels 0; Gaps 0;


```

Db 2162 AGTCACACTGTAGAGCTTCCCTGTGGCTTACCCACACACATCTCCCTCATCACTA 2221
Oy 2225 ccttgcctgaagctactctggtaagagcttgcgaagaatacagaaggaatctg 2284
Db 2222 CCTTGCTGAGCTACTGCTGAGTAAGAGACGCCAGAGACTATATATAAAGCCAAATGTG 2281
Oy 2285 tcacattcaagctctctcctcaactctgtctctgattcgactctctctcacaagcgccacg 2344
Db 2282 TCACCTTACGCTGCTGCTCTCAACTTCATCTGATGAGCTCTTCTTCAACATGCGCACCA 2341
Oy 2345 tctcagcagcaaglaactctctcgcgcgaacatgatagtctgagctgagcagcctgagca 2404
Db 2342 TTTTACAGAGCAGCTACTGCTGCTGCGGTCAATGTCTGCGACAGGCGCTGACACACTGAGCG 2401
Oy 2405 ggcgcttcggtggttattcttctgcttaagtctgaagtactctctgcgcgcgaagccca 2464
Db 2402 GCGGCTTACGCGGTACTTCTCTCCCAAGTGTATGTATCTCTGCGCTGACAGACTCA 2461
Oy 2465 acaagcagagactctccagagcctccatctcaggaactacagagcgctcgctccacc 2523
Db 2462 ACAATACAGAACACTTTCAGGCTCCATCCAGACTACAGAGCGCGCTGCGCACTAC 2520

RESULT 10
AX282896 1101 bp DNA linear PAT 02-NOV-2001
LOCUS Sequence 5 from Patent WO0172842.
DEFINITION AX282896
ACCESSION AX282896
VERSION AX282896.1 GI:16609870
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
AUTHORS Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,
Abdin,A., Zambrowicz,B. and Sands,A.T.
TITLE Novel human 7tm proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0172842-A 5 04-OCT-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
source 1..1101
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 219 a 308 c 323 g 250 t 1 others
ORIGIN
Query Match 39.6%; Score 1001.2; DB 6; Length 1101;
Best Local Similarity 99.6%; Pred. No. 3,1e-207;
Matches 1004; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Db 301 GCGTCAGAACGCTGGGCCCTCTCCAGGACATCACTGGGGTGGCCGGATCCAGGCAT 360
Oy 955 gggatgtgtctggtggtctggtccatccagaagaggtcttccctggtcgttaaggtcttaa 1014
Db 361 GGGATGTGTCTGGGGGTGGCCATCCAGAAAGGGCTGTCCCTGGGCTTAAGCGCTTTGAA 420
Oy 1015 gaagcctatgcccgggagagaagagcccttaagccttgctccaaagaggtcctgtgtgc 1074
Db 421 GAAAGCTTATGCCCGGGGAGAGAGAGGCCCTTAAAGCTTGGCCAGAGGCTCTGCTGTC 480
Oy 1075 agcagcaatcagctctcagagaatgccaagcttcatgacacacacagctgcccagctc 1134
Db 481 AGCAGCAATACAGCTCTGAGAGAAATGACAGCTTCATGRCACACAGATGCCAAGCTC 540
Oy 1135 aaagcctctcagatgagctctgctcacaagcagatcagcggtctgtataggtgtgccc 1194
Db 541 AAAGCCTTCTCCATGATGTTGTGCTTACAAAGCATCCGCGCTGTGTATGCGGTGGCCAT 600
Oy 1195 ggcctccacagctctcgtggtctgtcctctgagcttgttccaggggcagagctacc 1254
Db 601 GGCTTCCACAGCTCTCTGGGCTGTGCTCTGTGAGCTTGTTCAGAGGCGCGAGCTTACCC 660
Oy 1255 tggcagctcttggagcagatccacaaggttgatctctctcacaagaagacgtgtggcg 1314
Db 661 TGGCAGCTTTTGGAGCAGATCCACAAGGTGATTTCTTTCACACAGAGACAGCTGGCG 720
Oy 1315 tttaatgcaacaagagatccctcagtagtgcatacaaatgtctgagctgtgagaaatga 1374
Db 721 TTTAATGACAAACAGAGATCCCTCAGTAGCTATAAATAATTTGCTGGAGCTGGAATGA 780
Oy 1375 cccaagtgagcttcaagctcgtctcgtctccacacatgctcagctcagctaaacata 1434
Db 781 CCCAGTGGAGCTTACCGGTCTCGTGTCTTCCATATGCTCCAGTTCAGCTAATACATA 840
Oy 1435 aatgagacaaatccacagtgagcaggaaggaacacagagtgctaaatctgtgtctcc 1494
Db 841 AATGAGACCAAAATCACTGAGCAGCAAGAAAGACACAGAGTCCATATGCTGTGTTCC 900
Oy 1495 agcagctcttgaagggcagcagagatgttgaaggtttccatcatctgcgtttgag 1554
Db 901 AGCGACTGTCTTGAAGGCGACAGCGAGTGTACGGGTTCATCATCTGCTGCTTGA 960
Oy 1555 tgtgtgacctgtgggctgtgagccttctcacaagaagtgacctctac 1602
Db 961 TGTGTGCCCTGTGGGCTGGAGCTTCTCAACAGAGTATATCTCTAC 1008

RESULT 11
AL591866 102539 bp DNA linear PRI 25-OCT-2001
LOCUS Human DNA sequence from clone Rpl11-58A11 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL591866
VERSION AL591866.13 GI:16501194
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 102539)
AUTHORS Wallis,J.
JOURNAL Direct Submission
Submitted (25-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clone@sanger.ac.uk
On Oct 26, 2001 this sequence version replaced gi:15022299.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
only a small overlap as described above.
COMMENT

```

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C-elegans/wormpep>. This sequence was generated from part of bacterial clone configs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>. Rpl1-58A11 is from the library RPl1-11.1 constructed by the group of Pletier de Jong. For further details see <http://www.chori.org/bacpac/home.htm>. VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RPl1-58A11. It may be shorter because we sequence overlapping sections only once, except for a short overlap. The true right end of clone RPl1-58A11 is at 102539 in this sequence. The true left end of clone RPl1-239P22 is at 13247 in this sequence. The true right end of clone RPl1-650H14 is at 2000 in this sequence.

FEATURES

source 1..102539
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RPl1-58A11"
/clone_lib="RPl1-11.1"
582..618
/note="Sequence from overlapping clone RPl1-650H14 (AL158217). Assembly confirmed by restriction digest."
1452..1473
/note="Sequence from overlapping clone RPl1-650H14 (AL158217). Assembly confirmed by restriction digest."
48541..48603
/note="Sequence from overlapping clone RPl1-239P22 (AL159177). Assembly confirmed by restriction digest."
69007..69267
/note="Sequence from overlapping clone RPl1-239P22 (AL159177). Assembly confirmed by restriction digest."
100154..100179
/note="Sequence from overlapping clone RPl1-239P22 (AL159177). Assembly confirmed by restriction digest."
misc_feature 25289 a 26768 c 27360 g 23122 t
BASE COUNT 25289 a 26768 c 27360 g 23122 t
ORIGIN

Query Match 36.9%; Score 932.6; DB 9; Length 102539;
Best Local Similarity 99.9%; Pred. No. 2.6e-192;

Matches 932: Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1594 gacgtctacagatgacgactgtgtggaagaagatgggacactgaggaacagacc 1653
|||||
DB 75842 GACGCTACAGATGCCACTGTGTGGAAAGAGTGGGACACTGAGGAACCCAGACC 75901
QY 1654 tgcctccgcgcacgtgtgtgtttgtgcttgcgtgagcacacgtctgtgtgtcgtc 1713
|||||
DB 75902 TCGTTCGCCGCGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 75961
QY 1714 gtagctaaacgt 1773
|||||
DB 75962 GGAGCTAAACGCTGT 76021
QY 1774 ctgagacccctgt 1833
|||||
DB 76022 CTAGAACCCCTGT 76081

QY 1834 ctgagcagcagtagtagtggcagcctctatgtctctttgtgggaaccacagcctgtcgtc 1893
|||||
DB 76082 CTGGCAGCAGTAGTAGTGGCAGCCCTGTATGGCTTCTTGGGGAACCCACAGGCCCTGGTGC 76141
QY 1894 ttgtacgcagcagcctcttttgccttggtttccacatcttctgtctcgtctgaagt 1953
|||||
DB 76142 TTGCTATGCGCAGGCGCCCTTGTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 76201
QY 1954 cgtcaatcacaactaatcaatcaatcttcaagtttccacagaagttacatctatccac 2013
|||||
DB 76202 CGCTATTCCTCACTATATCATCTTCAAGTTTCCACCAAGTACTATCATCTTCAACAC 76261
QY 2014 gctctgttccaaacacacagcgttgcgtctgttgtgtgtgtgtgtgtgtgtgtgtgtgt 2073
|||||
DB 76262 GCGTGGGTCCAAACACACAGGT 76321
QY 2074 ctatctgtctaatctgt 2133
|||||
DB 76322 CTATCTGTCTAACTGT 76381
QY 2134 ttcccccactgt 2193
|||||
DB 76382 TTCCCCATCTGT 76441
QY 2194 ttctctacaatgtgcctctctctccatcagtgcttgccttgcctgaagcctgtgtgtgtgt 2253
|||||
DB 76442 TTCTCTTCAATGAGCT 76501
QY 2254 ttgcagagaactacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2313
|||||
DB 76502 TTGCCAGAACTACACAGCAGGCGCAATGTGTACCTTGAAGCTGTCTTCACTTGTGTGT 76561
QY 2314 tcttgatgcctctctctctccacagcagcagcagcagcagcagcagcagcagcagcagcagc 2373
|||||
DB 76562 TCTGTGATGCGCTTCTTTCACACAGGCGCAGCTGTACAGCGCAAGTACTGCTGGCGCC 76621
QY 2374 aacatgt 2433
|||||
DB 76622 AACATATGCTGTGGGT 76681
QY 2434 tcttaagtatctctctgtgc 2493
|||||
DB 76682 TGTACTGATCTCTTGTGCGGCCGACAGCTCTACAGCAGCAGCAGCAGCAGCAGCAGCAG 76741
QY 2494 caggactacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2526
|||||
DB 76742 CAGGACTACACAGGCGGCTGCGGCTCCACCTGA 76774
RESULT 12
AC108123 191481 bp DNA linear HTG 25-JAN-2002
LOCUS AC108123/c
DEFINITION Homo sapiens chromosome 16 clone RPl1-56511, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
ACCESSION AC108123.1 GI:18369969
VERSION AC108123.1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 191481)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 191481)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (25-JAN-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI

RESULT 15
 AL611927/c 198306 bp DNA linear HTG 30-JAN-2002
 LOCUS Mus musculus chromosome 4 clone RP23-445E20, *** SEQUENCING IN
 DEFINITION PROGRESS ***, in unordered pieces.
 ACCESSION AL611927 GI:18476802
 VERSION HTG5_PHASE1; HTG5_ACTIVEFIN; HTG5_DRAFT; HTG5_FULLTOP.
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (sites)
 AUTHORS Clark, G.
 JOURNAL Direct Submission
 Submitted (23-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On Feb 1, 2002 this sequence version replaced gi:18135147.
 COMMENT ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 Project Information
 Center project name: BM445E20
 ----- Summary Statistics
 Sequencing vector: plasmid: L08752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Consensus quality: 197720 bases at least Q40
 Consensus quality: 197865 bases at least Q30
 Consensus quality: 197958 bases at least Q20
 Insert size: 198106; sum-of-contigs
 Quality coverage: 207675; 0.3% error; agarose-fp
 Quality coverage: 8.45x in Q20 bases; sum-of-contigs Quality
 coverage: 8.06x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES

source

1. 198306
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="4"
 /clone="RP23-445E20"
 /clone_11b="RPC1-23"
 1. 83829
 /note="assembly_fragment:03510
 fragment_chain:1"
 misc_feature
 83930..174680
 /note="assembly_fragment:03467
 fragment_chain:1"
 misc_feature
 174781..198306
 /note="assembly_fragment:05008
 fragment_chain:1"

BASE COUNT 47809 a 49338 c 52220 g 48737 t 202 others
 ORIGIN

Query Match 25.9%; Score 654.2; DB 2; Length 198306;
 Best Local Similarity 81.2%; Pred. No. 8.9e-132;
 Matches 758; Conservative 1; Mismatches 174; Indels 0; Gaps 0;

QY 1594 gacctctacagatgcacgtctgtggaagaagatggcactgaaggaagcagacc 1653
 11
 Db 116960 GAGCTTCACACGCTGCGCCTGTGGAACAGAAATGGGCCCTCAGGGAGCTCAGCC 116901
 QY 1654 tgcctccgcgcacgtgtgtttttgtgcttgcgtgagcaacacctctgtgtcgtcgtg 1713
 11

Db 116900 TGGTCTCAACGACCGTGGAGTTCTTGGGGTGGCATGAACCCATCTCTTGGTCTATTAA 116841
 QY 1714 gcaactaacacgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtc 1773
 11
 Db 116840 GACGCTACACGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 116781
 QY 1774 cttagacacccctgtc 1833
 11
 Db 116780 CTTCACACCCCTGTTGTGAGTGCAGCTGGGGAGCGCTGTCTCTCAATGCTGGGATTCC 116721
 QY 1834 ctggcagcaggtc 1893
 11
 Db 116720 TTGGTACGTGGAGTGTGCAACCTCTACAGCTTCTTGGAGACCCAGGTGCTGGCTGAC 116661
 QY 1894 ttgctaacgcagcgcctctctgt 1953
 11
 Db 116660 TTGCTCGTACAGCCCTCTTCTTCTCGGGATTTCATTTCTCTCTCTCTCTCTCTCTCT 116601
 QY 1954 cgcctactccaactaatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatc 2013
 11
 Db 116600 CGCTCTTCCAACTGGTGCATCATCTTCAAGTTTTCACCAAGGTACCCACATTTCAACAC 116541
 QY 2014 gcttggtctcaaaacacagcgt 2073
 11
 Db 116540 ACTTGGGCCCAAAACCATGAGTGGCGGAATATTCGATTTGTCACTCCACGGCTCAATTGG 116481
 QY 2074 ctatctgtctaaactgtc 2133
 11
 Db 116480 TTGCTGTGCTACGTTGGGCTTGGCAATGTGAGCCGCCACCGGCGCCACGAGTTCAGGCC 116421
 QY 2134 ttcccccactgtc 2193
 11
 Db 116420 TTCCCCCACTGTGTGATTTCTTGAAGTGCACAGAGTCACTGTGTGGCTTCTGTGGTGT 116361
 QY 2194 ttcccttaaatgtc 2253
 11
 Db 116360 TTGGCACACACATCT 116301
 QY 2254 ttgccagagaactaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2313
 11
 Db 116300 CTGCGGAGACATATACAGAGCAATGTGTACCTTGACCTTGACCTTGACCTTGACCTTG 116241
 QY 2314 tccctgactgcctctctctcaacacgcagcagcagcagcagcagcagcagcagcagcagcagc 2373
 11
 Db 116240 TCTGTGATGCTTTCTTTCACACATGTCCACGATTTACAGGGCGAGCTACCCCGGGCTC 116181
 QY 2374 aacatgatgt 2433
 11
 Db 116180 AATGTCTGTGGAGGCTGTGGCTGTGAGTGGGCTTTCAGCGCTATTTCTCTCTCTTAA 116121
 QY 2434 tgcctacgtatcctctgtccgcagcagcagcagcagcagcagcagcagcagcagcagcagc 2493
 11
 Db 116120 TGTGTAGTATTTCTGTGCGGCTCCAGAACTCAACACAGAACTTTCAGGCTTCATC 116061
 QY 2494 caaggactaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2526
 11
 Db 116060 CAGGACTACACGAGCGCTGCGGCACTACCTGA 116028

Search completed: June 11, 2002, 06:57:24
 Job time: 1165 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 11, 2002, 06:48:18 ; Search time 1949.94 Seconds
(without alignments)
17484.286 Million cell updates/sec

Title: US-09-819-946-1
Perfect score: 2526
Sequence: 1 atgtctgtctgtcagcgtctg.....ggcgtctgtcgtccacctga 2526

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 13736207 segs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	369	14.6	938	10	BI462667 603202107
2	299	11.8	562	9	AI742401 wq40e02.x
3	140	5.5	496	9	AA853967 a51e10.s
4	66	2.6	120	12	BH173277 0011857 m
5	43	1.7	514	12	AO615201 HS 5138 B
6	27	1.1	193	12	AO631494 RCT-11-4
7	26	1.0	473	10	BI580014 RE74133.5
8	26	1.0	522	10	BI485890 RE69243.5
9	26	1.0	522	12	BH093096 RPCI-24-3
10	26	1.0	536	10	BI213517 R18934.5
11	26	1.0	536	10	BI235687 RE31596.5
12	26	1.0	543	12	AZ816179 2M0084113
13	26	1.0	585	10	BI577858 RE71616.5
14	26	1.0	586	10	BI366613 RE52188.5
15	26	1.0	611	10	BF500786 AV15818.5
16	26	1.0	648	10	BI368783 RE54819.5
17	26	1.0	667	9	AI546359 LD47919.5

C 18	26	1.0	685	10	BI577296
C 19	26	1.0	709	10	BF505332
C 20	26	1.0	712	12	BH385657
C 21	25	1.0	168	10	BF421445
C 22	25	1.0	339	9	AI759574
C 23	25	1.0	407	12	AZ699387
C 24	25	1.0	501	10	BI895264
C 25	25	1.0	559	12	BH389069
C 26	25	1.0	564	12	BH376112
C 27	25	1.0	564	12	FR0008457
C 28	25	1.0	652	9	BB664716
C 29	25	1.0	654	10	BF123370
C 30	25	1.0	676	9	AL507968
C 31	25	1.0	687	9	BB610646
C 32	25	1.0	699	10	BF168171
C 33	25	1.0	700	9	AL503561
C 34	25	1.0	760	10	BI647065
C 35	25	1.0	765	10	BI557461
C 36	25	1.0	783	12	BH375795
C 37	25	1.0	789	10	BG861598
C 38	25	1.0	791	10	BI410778
C 39	25	1.0	808	10	BE412655
C 40	25	1.0	819	10	BG919691
C 41	25	1.0	853	10	BF631482
C 42	25	1.0	888	10	BI690643
C 43	25	1.0	925	12	CNS0381E
C 44	25	1.0	961	12	CNS053CV
C 45	25	1.0	1017	10	BM465643

ALIGNMENTS

RESULT 1
BI462667 938 bp mRNA linear EST 21-AUG-2001
LOCUS 603202107F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5268264 5',
DEFINITION mRNA sequence.
ACCESSION BI462667.1 GI:15253323
VERSION BI462667.1
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS 1 (bases 1 to 938)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: ggaaps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LHAM1676 row: m column: 01
High quality sequence stop: 739.
Location/Qualifiers
1. 938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5268264"
/clone_lid="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag) ; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched

FEATURES

Source

Query Match 1.7%; Score 43; DB 12; Length 514;
 Best Local Similarity 100.0%; Pred. No. 6.9e-10;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1543 tctgtcttgagtggtgcccctgaggaacctctca 1585
 |||||||
 Db 280 TGCTGCTTGAGTGTGCTGCGGCGGACCTTCTCA 238

RESULT 6
 A0631494/c 193 bp DNA linear GSS 17-JUN-1999
 DEFINITION RPCI-11-473E18.TV RPCI-11 Homo sapiens genomic clone RPCI-11-473E18
 , DNA sequence.
 ACCESSION A0631494
 VERSION A0631494.1 GI:5094129
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 193)
 Zhao, S., Adams, M.D., Nieman, W., Malek, J., de Jong, P. and Venter
 , J.C.
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 Unpublished (1997)
 Contact: Shaying Zhao, William Nieman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbeetigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pletier de Jong
 (pletier@edj.med.bufileo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.bufileo.edu/ordering>) or from
 Research Genet cs (<http://inforesgen.com>). BAC end search page:
http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html.
 Seq primer: 17
 Class: BAC ends.
 Location/Qualifiers
 1. 193
 /organism="Homo sapiens"
 /db_xref="GDB:7681361"
 /db_xref="taxon:9606"
 /clone="RPCI-11-473E18"
 /clone_1lb="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPCI11 Human Male BAC Library"

BASE COUNT 36 a 66 c 52 g 39 t
 ORIGIN

Query Match 1.1%; Score 27; DB 12; Length 193;
 Best Local Similarity 100.0%; Pred. No. 0.046;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1543 tctgtcttgagtggtgcccctgaggaacctctca 1569
 |||||||
 Db 193 TGCTGCTTGAGTGTGCTGCGGCGGACCTTCTCA 167

RESULT 7
 B1580014/c 473 bp mRNA linear EST 06-SEP-2001
 LOCUS B1580014
 DEFINITION R74133.5prime RE Drosophila melanogaster normalized Embryo pflc-1
 Drosophila melanogaster cDNA clone R74133 5 similar to CG18543:

FBan0018543 GO:[] located on: 3L 66C12-66C12.: 05/21/2001. mRNA
 sequence.
 B1580014
 VERSION B1580014.1 GI:15471436
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 473)
 Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
 , J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
 , R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mistra, S.,
 Mungall, C.J., Nuno, J., Pacle, J., Paragas, V., Park, S.,
 Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin
 , G.M.
 BDGP/HMI RE Drosophila EST project
 Unpublished (2001)
 Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_estfruitfly.berkeley.edu
 hit genomic AE003555: arm:3L [8254613,8536879]
 estimated-cyto:66C8-66D9: 05/21/2001
 Plate: RE.741 row: C column: 9
 High quality sequence stop: 449.
 Location/Qualifiers
 1. 473
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="R74133"
 /clone_1lb="RE Drosophila melanogaster normalized Embryo
 pflc-1"
 /sex="male and female"
 /dev_stage="0-24 hours mixed stage embryonic"
 /lab_host="DH5-alpha Tona"
 /note="Organ: embryo; Vector: pBlc1; Site_1: XhoI; Site_2:
 BamHI; Library was kindly generated by Piero Carinai at
 the RIKEN. The library was normalized and excised using
 Cre recombinase. Plasmid cDNA library."

BASE COUNT 135 a 140 c 111 g 87 t
 ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 473;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1725 gctgctgctgctgctgctgctgga 1750
 |||||||
 Db 328 GCTGCTGCTGCTGCTGCTGCTGGA 303

RESULT 8
 B1485890 522 bp mRNA linear EST 28-AUG-2001
 LOCUS B1485890/c
 DEFINITION RE69243.5prime RE Drosophila melanogaster normalized Embryo pflc-1
 Drosophila melanogaster cDNA clone RE69243 5 similar to CG18543:
 FBan0018543 GO:[] located on: 3L 66C12-66C12.: 05/16/2001. mRNA
 sequence.
 ACCESSION B1485890
 VERSION B1485890.1 GI:15325673
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 522)
 Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson

J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C.J., Nunoo, J., Pacled, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin, G.M.

TITLE BDGP/HIMI RE Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT Contact: Stapleton, M.

BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
hit genomic: AE003555: arm:3L [8254613,8536879]
estimated-cyto: 66C8-66D9: 05/16/2001
Plate: RE.692 row: D column: 7
High quality sequence stop: 509.

FEATURES

source

1..522
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="RE69243"
/clone_lib="RE Drosophila melanogaster normalized Embryo

pflc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/note="Organ: embryo; Vector: pFlc1, Site_1: XhoI, Site_2: BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

BASE COUNT

146 a 155 c 123 g 98 t

ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 522;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgctgctgctgctgctgctgga 1750
|||||

Db 328 gctgctgctgctgctgctgctgga 303
|||||

RESULT 9

BH093096/c

LOCUS BH093096 522 bp DNA linear GSS 18-JUL-2001
DEFINITION RPCI-24-357D16.TV RPCI-24 Mus musculus genomic clone RPCI-24-357D16
, DNA sequence.

ACCESSION

BH093096

VERSION

BH093096.1

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Mammalia; Euthera; Rodentia; Sciurognath; Muridae; Murinae; Mus.

AUTHORS

Zhao, S., Niemman, W., Malek, J., Shatsman, S., Akincet, B., Levins, M.,

Russell, D., de Jong, P. and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other GSSs: RPCI-24-357D16.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC

Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end
page: http://www.tigr.org/db/bac_ends/mouse/bac_end_intro.html
Plate: 357 row: D column: 16
Seq primer: 77
Class: BAC ends.

FEATURES
source

1..522
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-357D16"
/clone_lib="RPCI-24"
/sex="Male"

/cell_type="Spleen/Brain"
/note="Vector: pIRBAC1; Site_1: BamHI; Site_2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pIRBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

BASE COUNT

152 a 137 c 120 g 113 t

ORIGIN

Query Match 1.0%; Score 26; DB 12; Length 522;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgctgctgctgctgctgctgga 1750
|||||

Db 141 gctgctgctgctgctgctgctgga 116
|||||

RESULT 10

BI213517/c

LOCUS BI213517 536 bp mRNA linear EST 11-JUL-2001
DEFINITION RE18934.Sprline RE Drosophila melanogaster normalized Embryo pflc-1
Drosophila melanogaster cDNA clone RE18934 5 similar to CG18543;
Fban0018543 'cell cycle regulator' located on: 3L 66C12-66C12.1;
04/12/2001, mRNA sequence.

ACCESSION

BI213517

VERSION

BI213517.1

KEYWORDS

EST.

SOURCE

fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 536)

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson

J., Champagne, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S.,

Mungall, C.J., Nunoo, J., Pacled, J., Paragas, V., Park, S.,

Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin

G.M.

BDGP/HIMI RE Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.

BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
hit genomic: AE003555: arm:3L [8254613,8536879]
estimated-cyto: 66C8-66D9: 04/12/2001
Plate: RE.189 row: C column: 10
High quality sequence stop: 524.

FEATURES

1..536
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="RE18934"
/clone_lib="RE Drosophila melanogaster normalized Embryo

pflc-1"

/sex="male and female"
 /dev_stage="0-24 hours mixed stage embryonic"
 /lab_host="DH5-alpha Tona"
 /note="Organ: embryo; Vector: pF1c1; Site: 1: XhoI; Site: 2: BamHI; Library was kindly generated by Piero Carinini at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

BASE COUNT 146 a 158 c 128 g 104 t
 ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 536;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgcctcgtcgtcgtcgttggga 1750
 ||||||||||||||||||||||||||||
 Db 329 GCTGCTGCTGCTGCTGCTGCTGGA 304

RESULT 11
 B1235687 536 bp mRNA linear EST 12-JUL-2001
 LOCUS B1235687/c
 DEFINITION RE31596.5prime RE Drosophila melanogaster normalized Embryo pF1c-1
 Drosophila melanogaster cDNA clone RE31596 5 similar to CG18543:
 FBan0018543 'cell cycle regulator' located on: 3L 66C12-66C12.1;
 05/12/2001, mRNA sequence.

ACCESSION B1235687
 VERSION B1235687.1 GI:14704130
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 536)
 Stapleton, M., Brkstein, P., Hong, L., Tyler, D., Berman, B., Carlson
 J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
 R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S.,
 Munhall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,
 Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celinker, S. and Rubin
 G. M.

TITLE BDF/HMMI RE Drosophila EST project
 JOURNAL Unpublished (2001)
 COMMENT BDCP
 Contact: Stapleton, M.
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
 hit genomic A6003555; arm: 3L [8254613, 8536879]
 estimated-cyto: 66C8-66D9: 05/12/2001
 Plate: RE.315 row: H column: 12
 High quality sequence stop: 494.

FEATURES
 source

1..536
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_1ib="RE31596"
 /clone_1ib="RE Drosophila melanogaster normalized Embryo
 pF1c-1"
 /sex="male and female"
 /dev_stage="0-24 hours mixed stage embryonic"
 /lab_host="DH5-alpha Tona"
 /note="Organ: embryo; Vector: pF1c1; Site: 1: XhoI; Site: 2:
 BamHI; Library was kindly generated by Piero Carinini at
 the RIKEN. The library was normalized and excised using
 Cre recombinase. Plasmid cDNA library."

BASE COUNT 148 a 157 c 128 g 103 t
 ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 536;

Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgcctcgtcgtcgtcgttggga 1750
 ||||||||||||||||||||||||||||
 Db 329 GCTGCTGCTGCTGCTGCTGCTGGA 304

RESULT 12
 A2816179 543 bp DNA linear GSS 20-FEB-2001
 LOCUS A2816179/c
 DEFINITION 2M0084L13R Mouse 10kb plasmid UGCC1M library Mus musculus genomic
 clone UGCC2M0084L13 R, DNA sequence.

ACCESSION A2816179
 VERSION A2816179.1 GI:12986087
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 543)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
 and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
 Insert Length: 1000 Std Error: 0.00
 Plate: 0084 row: L column: 13
 Seq primer: CACACGAGAAACAGCATATGACC
 Class: plasmid ends
 High quality sequence stop: 543.

FEATURES
 source

1..543
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_1ib="UGCC2M0084L13"
 /clone_1ib="Mouse 10kb plasmid UGCC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42ny; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (<http://www.jax.org/resources/documents/dnares/>). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD42 (g11473211419b/AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 152 a 139 c 105 g 147 t
 ORIGIN

Query Match 1.0%; Score 26; DB 12; Length 543;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgctgctgctgctgctgctgga 1750
 ||||||||||||||||||||||||||||
 Db 197 gctgctgctgctgctgctgctgga 172

RESULT 13
 B1577858/c
 LOCUS 585 bp mRNA linear EST 06-SEP-2001
 DEFINITION RE1616.5prime RE Drosophila melanogaster normalized Embryo pf1c-1
 Drosophila melanogaster cDNA clone RE1616 5 similar to CG18543:
 Fban0018543 GO:[] located on: 3L 66C12-66C12; 05/17/2001, mRNA
 sequence.

ACCESSION B1577858
 VERSION B1577858
 KEYWORDS GI:15469280
 SOURCE EST.
 ORGANISM fruit fly.
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 585)
 AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
 ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
 ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mista,S.,
 Mungall,C.J., Nunoo,J., Pacled,J., Paragas,V., Park,S.,
 Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
 ,G.M.
 BDGP/HMI RE Drosophila EST Project
 TITLE B1577858 RE Drosophila EST Project
 JOURNAL Unpublished (2001)
 COMMENT Contact: Stapleton, M.
 BDCP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 hit genomic AE003555: arm:3L [8254613,8536879]
 estimated-cyto:66C8-66D9: 05/17/2001
 Plate: RE:716 row: B column: 4
 Location/Qualifiers
 1..585
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_id="RE1616"
 /clone_1lb="RE Drosophila melanogaster normalized Embryo
 pf1c-1"
 /sex="male and female"
 /dev_stage="0-24 hours mixed stage embryonic"
 /lab_host="DH5-alpha Tona"
 /note="Organ: embryo; Vector: pf1c1; Site_1: XhoI; Site_2:
 BamHI; Library was kindly generated by Piero Carninci at
 the RIKEN. The library was normalized and excised using
 Cre recombinase. Plasmid cDNA library."

BASE COUNT 162 a 169 c 141 g 113 t

ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 585;
 Best Local Similarity 100.0%; Pred. No. 0.19;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgctgctgctgctgctgctgga 1750
 ||||||||||||||||||||||||||||
 Db 328 gctgctgctgctgctgctgctgga 303

RESULT 14
 B1366613/c
 LOCUS 586 bp mRNA linear EST 01-AUG-2001

DEFINITION RE52188.5prime RE Drosophila melanogaster normalized Embryo pf1c-1
 Drosophila melanogaster cDNA clone RE52188 5 similar to CG18543:
 Fban0018543 'cell cycle regulator' located on: 3L 66C12-66C12;:
 05/14/2001, mRNA sequence.

ACCESSION B1366613
 VERSION B1366613
 KEYWORDS GI:15062641
 SOURCE EST.
 ORGANISM fruit fly.
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 586)
 AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
 ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
 ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mista,S.,
 Mungall,C.J., Nunoo,J., Pacled,J., Paragas,V., Park,S.,
 Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
 ,G.M.
 BDGP/HMI RE Drosophila EST Project
 TITLE B1366613 RE Drosophila EST Project
 JOURNAL Unpublished (2001)
 COMMENT Contact: Stapleton, M.
 BDCP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 hit genomic AE003555: arm:3L [8254613,8536879]
 estimated-cyto:66C8-66D9: 05/14/2001
 Plate: RE:521 row: H column: 4
 Location/Qualifiers
 1..586
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_id="RE52188"
 /clone_1lb="RE Drosophila melanogaster normalized Embryo
 pf1c-1"
 /sex="male and female"
 /dev_stage="0-24 hours mixed stage embryonic"
 /lab_host="DH5-alpha Tona"
 /note="Organ: embryo; Vector: pf1c1; Site_1: XhoI; Site_2:
 BamHI; Library was kindly generated by Piero Carninci at
 the RIKEN. The library was normalized and excised using
 Cre recombinase. Plasmid cDNA library."

BASE COUNT 162 a 169 c 142 g 113 t

ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 586;
 Best Local Similarity 100.0%; Pred. No. 0.19;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgctgctgctgctgctgctgga 1750
 ||||||||||||||||||||||||||||
 Db 328 gctgctgctgctgctgctgctgga 303

RESULT 15
 Bf500786/c
 LOCUS 611 bp mRNA linear EST 19-APR-2001
 DEFINITION AT15818.5prime AT Drosophila melanogaster adult testes pOTB7
 Drosophila melanogaster cDNA clone AT15818 5 similar to CG13798:
 Fban0013798 located on: 3L 62E4-62E4; 04/09/2001, mRNA sequence.

ACCESSION Bf500786
 VERSION Bf500786
 KEYWORDS GI:13692625
 SOURCE EST.
 ORGANISM fruit fly.
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 611)

AUTHORS

Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E., Berman, B., Carlson, J., Champagne, M., Chavez, C., Chew, M., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phoumenavong, S., Wan, K., Yu, C., Lewis, S. E., Celisner, S., and Rubin, G. M.

TITLE

JOURNAL

BDGP/HHMI AT Drosophila EST Project
Unpublished (2000)

COMMENT

On Dec 6, 2000 this sequence version replaced gl:11584087.
Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AEO03474: arm:3r [2120287,2430124]

estimated-cyto:62C4-62E5: 04/09/2001

Place: AT:158 row: B column: 6

High quality sequence stop: 571.

FEATURES

source

1. 611
Location/Qualifiers

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone="AT15818"

/clone_lib="AT Drosophila melanogaster adult testes potB7"

/sex="male"

/dev_stage="0-3 day old Ore-R males"

/lab_host="Plates AT.10-AT.120: DHS-alpha. Plates

AT.121-AT.319: DHS-alpha Tona"

/note="Organ: ADULT testes; Vector: potB7; Site_1: EcORI;

Site_2: XhoI; The mRNA for the testis library was made

from testes and seminal vesicles hand dissected from 0-3

day old Ore-R males. RNA kindly provided by the lab of

Margaret Fuller. Sized fractionated cDNAs were directly

ligated into potB7. Plasmid cDNA library."

BASE COUNT

169 a 190 c 164 g 88 t

ORIGIN

Query Match

1.0%; Score 26; DB 10; Length 611;

Best Local Similarity 100.0%; Pred. No. 0.19;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1725 gctgctgctgctgctgctgctgctgga 1750

|||||

DB 345 GCTGCTGCTGCTGCTGCTGCTGGA 320

Search completed: June 11, 2002, 08:20:22
Job time: 5524 sec

OTHER INFORMATION: repeat sequence"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 104..252
OTHER INFORMATION: /note="Nucleotides 104-252 are
OTHER INFORMATION: unique flanking sequence"
US-08-623-906A-1

Query Match 0.9%; Score 23; DB 2; Length 252;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1724 ccgtgctgctgctgctgctgctt 1746
|||||
Db 27 ccgtgctgctgctgctgctgctt 49

RESULT 2
US-08-863-639A-31/c
; Sequence 31, Application US/08863639A
; Patent No. 5981185
; GENERAL INFORMATION:
; APPLICANT: Matson, Robert S.
; APPLICANT: Coassin, Peter J.
; APPLICANT: Rampal, Jang B.
; APPLICANT: Caskey, C. T.
; TITLE OF INVENTION: OLIGONUCLEOTIDE REPEAT ARRAYS
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, 9th Floor
; CITY: Pasadena
; STATE: CA
; COUNTRY: USA
; ZIP: 91101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Corel Wordperfect 8 version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/863,639A
; FILING DATE: May 28, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Joseph E. Muehl
; REGISTRATION NUMBER: 20,532
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 796-4000
; TELEFAX: (626) 795-6321
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid
US-08-863-639A-31

Query Match 0.9%; Score 22; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1724 ccgtgctgctgctgctgctgct 1745
|||||
Db 27 ccgtgctgctgctgctgctgct 6

RESULT 3
US-09-253-691-3/c

; Sequence 3, Application US/09253691
; Patent No. 6124100
; GENERAL INFORMATION:
; APPLICANT: Dong Kyu JIN
; TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases
; FILE REFERENCE: 1942/36
; CURRENT APPLICATION NUMBER: US/09/253,691
; CURRENT FILING DATE: 1999-02-22
; EARLIER APPLICATION NUMBER: KR 98-6,278
; EARLIER FILING DATE: 1996-02-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Wordperfect 6.1/Windows
; SEQ ID NO 3
; LENGTH: 397
; TYPE: DNA
; ORGANISM: human
US-09-253-691-3

Query Match 0.9%; Score 22; DB 3; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1724 ccgtgctgctgctgctgctgct 1745
|||||
Db 357 ccgtgctgctgctgctgctgct 336

RESULT 4
US-08-831-399-3
; Sequence 3, Application US/08831399
; Patent No. 6312916
; GENERAL INFORMATION:
; APPLICANT: Kopetzki, Erhard; Muller, Rainer;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core
; NUMBER OF SEQUENCES: 16
; TITLE OF INVENTION: Streptavidin Mutants
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831,399
; FILING DATE: 1-April-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6312916man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 604 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 44..499
OTHER INFORMATION:
OTHER INFORMATION: Positions 44..115 correspond to sig
peptide and 116..499 to mat peptide
US-08-831-399-3

Query Match 0.9%; Score 22; DB 4; Length 604;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1724 cgctgctgctgctgctgct 1745
|||||
DB 63 CGCTGCTGCTGCTGCTGCT 84

RESULT 5
US-09-228-986-38
Sequence 38, Application US/09228986
Patent No. 6359198
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Niels
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and their use in the Modification of Plant Cell Signalling
FILE REFERENCE: 11000/1020
CURRENT APPLICATION NUMBER: US/09/228.986
CURRENT FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 130
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 38
LENGTH: 822
TYPE: DNA
ORGANISM: Eucalyptus grandis
US-09-228-986-38

Query Match 0.9%; Score 22; DB 4; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1724 cgctgctgctgctgctgct 1745
|||||
DB 132 cgctgctgctgctgctgct 153

RESULT 6
US-08-240-124-1
Sequence 1, Application US/08240124
Patent No. 5516558
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240.124
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1037 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: hex-L A2
FEATURE:
NAME/KEY: CDS
LOCATION: 83..799
NAME/KEY: sig_peptide
LOCATION: 83..139
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 140..796
US-08-240-124-1

Query Match 0.9%; Score 22; DB 1; Length 1037;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1724 cgctgctgctgctgctgct 1745
|||||
DB 96 CGTGTGCTGCTGCTGCTGCT 117

RESULT 7
US-08-453-943-1
Sequence 1, Application US/08453943
Patent No. 5738844
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/453,943
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/240,124
FILING DATE: 09-MAY-1994
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1037 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: hek-L A2
FEATURE:
NAME/KEY: CDS
LOCATION: 83..799
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 83..139
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 140..796
US-08-453-943-1

Query Match 0.9%; Score 22; DB 1; Length 1037;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1724 cgctgctgctgctgctgct 1745
|||||
Db 96 CGCTGCTGCTGCTGCTGCT 117

RESULT 8
US-09-057-121-1
Sequence 1, Application US/09057121
Patent No. 5969110
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,121
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/240,124
FILING DATE:
APPLICATION NUMBER: US 08/161,132
FILING DATE:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1037 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: hek-L A2
FEATURE:
NAME/KEY: CDS
LOCATION: 83..799
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 83..139
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 140..796
US-09-057-121-1

Query Match 0.9%; Score 22; DB 2; Length 1037;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1724 cgctgctgctgctgctgct 1745
|||||
Db 96 CGCTGCTGCTGCTGCTGCT 117

RESULT 9
US-09-358-734-1
Sequence 1, Application US/09358734
Patent No. 6274117
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON

TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1776 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 36..1115
US-08-531-927B-1

Query Match 0.9%; Score 22; DB 2; Length 1776;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1724 cgctgctgctgctgctgct 1745
|||||

Db 988 CGCTGCTGCTGCTGCTGCT 967

RESULT 12
US-09-041-886-12/C
Sequence 12, Application US/09041886
Patent No. 6235872
GENERAL INFORMATION:
APPLICANT: Bredesen, Dale E.
TITLE OF INVENTION: Proapoptotic Peptides, Dependence
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1776 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 36..1116
US-09-041-886-12

Query Match 0.9%; Score 22; DB 4; Length 1776;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1724 cgctgctgctgctgctgct 1745

|||||
Db 988 CGCTGCTGCTGCTGCTGCT 967

RESULT 13
US-08-753-007A-5/C
Sequence 5, Application US/08753007A
Patent No. 6074841
GENERAL INFORMATION:
APPLICANT: Gearing, David P.
TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/753,007A
FILING DATE: 19-NOV-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/699,591
FILING DATE: 19-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07334/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1884 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 664...1883
OTHER INFORMATION:
US-08-753-007A-5

Query Match 0.9%; Score 22; DB 3; Length 1884;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1724 cgctgctgctgctgctgct 1745
|||||

Db 65 CGCTGCTGCTGCTGCTGCT 44

RESULT 14
US-09-398-496-5/C
Sequence 5, Application US/09398496
Patent No. 6133423
GENERAL INFORMATION:
APPLICANT: Gearing, David P.
TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
TITLE OF INVENTION: AND USES THEREFOR

NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/398,496
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/753,007
FILING DATE: 19-NOV-1996
APPLICATION NUMBER: 08/699,591
FILING DATE: 19-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07334/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1884 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 664..1883
OTHER INFORMATION:
US-09-398-496-5

Query Match 0.9%; Score 22; DB 3; Length 1884;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1724 cgcctgctgctgctgctgct 1745
|||||
DB 65 CGCTGCTGCTGCTGCTGCT 44

RESULT 15
US-08-261-822A-7/c
Sequence 7, Application US/08261822A
Patent No. 5650553
GENERAL INFORMATION:
APPLICANT: Ecker, Joseph R. et al.
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553rls
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,822A
FILING DATE: 17-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2234 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-261-822A-7

Query Match 0.9%; Score 22; DB 1; Length 2234;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1725 gctgctgctgctgctgctt 1746
|||||
DB 1970 GCTGCTGCTGCTGCTGCTT 1949

Search completed: June 11, 2002, 09:14:02
Job time: 8194 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: June 11, 2002, 06:49:22 ; Search time 3107.22 Seconds
(without alignments)
17012.130 Million cell updates/sec

Title: US-09-819-946-1
Perfect score: 2526
Sequence: 1 atgctgctctgcaagctcg.....ggcgctgcggtccactga 2526

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

word size :

Total. number of hits satisfying chosen parameters: 3595312

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Listing first 45 summaries

Database :

1:	gb_ba:	*
2:	gb_htg:	*
3:	gb_in:	*
4:	gb_om:	*
5:	gb_ov:	*
6:	gb_pat:	*
7:	gb_ph:	*
8:	gb_pl:	*
9:	gb_pr:	*
10:	gb_ro:	*
11:	gb_sts:	*
12:	gb_sy:	*
13:	gb_un:	*
14:	gb_vi:	*
15:	em_ba:	*
16:	em_fun:	*
17:	em_hum:	*
18:	em_in:	*
19:	em_mu:	*
20:	em_om:	*
21:	em_or:	*
22:	em_ov:	*
23:	em_pat:	*
24:	em_ph:	*
25:	em_pl:	*
26:	em_ro:	*
27:	em_sts:	*
28:	em_un:	*
29:	em_vi:	*
30:	em_htg_hum:	*
31:	em_htg_iny:	*
32:	em_htg_other:	*
33:	em_htgo_iny:	*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score	Match Length DB ID Description

[illegible]

ALIGNMENTS

[illegible]

REFERENCE

1 (sites)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

TITLE Novel human 7m proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0172842-A 1 04-OCT-2001;
Lexicon Genetics Incorporated (US)

source

BASE COUNT	469 a	770 c	697 g	587 t	3 others
ORIGIN					

Query Match	100.0%;	Score 2526;	DB 6;	Length 2526;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2526;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

QY	1021	latgcgcggagcagaacgaagagcccataaggccttgcacaaagggctcgtgtgcagcagc	108
Db	1021	TATGCCCCGGCAGACAAAGGAGGCGCCCTAGCGCTTGGCACAAAGGCGCTCGGTGCAGCAGC	108
QY	1081	aatacagcctcacaagaatgcacaagcttcaatgcacaagatgcacaagctcaaacgc	114
Db	1081	AATCAGCTCTGCACAGAAATGCCAAGCTTTATATRCACACAGATGCCAAGCTCAAAAGC	114
QY	1141	ttctccatgatgtctcgtccctcaacgcataacgagcctgtatgtcgtgtgcccataagcctc	120
Db	1141	TTTCTCATGAGTATCTGCGCTCAACAGCATACCGGGCGTGTATAGCGGTGACCATGGCCTC	120
QY	1201	caccagcctccctgggcctgtgcctcttgagagctgttccaggggcccaggtctcaaccctggcag	126
Db	1201	CACCAAGCTCCCGGGCTGTGCTGTGGAGCTGTTCACAGGGCGCGAGTCTACCCCTGGAG	126
QY	1261	cttttgagcagatccacaaggtgatcttccctctcaacaagaagacatgtgcygttat	132
Db	1261	CTTTTGGAGCAGATCCACAGGTGCATTTCTTACACAGACACTGTGGCGTTTAT	132
QY	1321	gacaaacagagatcccccagtaagctataataatgtcctggagactgaaatgagcccaag	138
Db	1321	GACAAACAGAGATCCCCCTACATAGCTATTAATATTGCTGGAGCTGAGATGAGACCCAG	138
QY	1381	tgagacctcaagagtcctccggttctcccaagatgtctccaggttcagctaaacataaataag	144
Db	1381	TGGACCTTACAGGTCCTCGGTCTCTCCACATGGTCTCCAGTTACAGTAAACATAATATAG	144
QY	1441	accaaaatccagatgagcaggaagaagaacaaacagatgcctaaatctgtgtgtccagcagc	150
Db	1441	ACCAAAATCCAGTGTGGCACGGGAAAGGACAAACAGATGCCATACTGTGTGTCCAGCGAC	150
QY	1501	tgctctgaaggagcaacagagatgttaacgggtttccatcactgtctgtcttgagatgtg	156
Db	1501	TGCTCTGAAGGACCAACGACAGTGTATACGGGTTTCCATCAGCTGCTTGTAGTGTGTG	156
QY	1561	ccctgttgaggcttggagaccttccctcaacaagaagtgaacctataagatgccaagccttgtg	162
Db	1561	CCCTGTGGGCGCTGGAGCCTTCCACAAAGAGGACCTCTACAGATGCCAGCCTTGTGGG	162
QY	1621	aaagaagaagtggagcaacttgaggggaagccagacactgtctccgcgcacgtgtgttttg	168
Db	1621	AAAGAAGAAGTGGGCACTGTAGGGGAACCCACCACTGTCTCCGCGACGTGTGTTTGG	168
QY	1681	gcttgcgtgagcaacacctcttggtgtgtcgtgcaagcacaagcctcgtcgtctgtcgtg	174
Db	1681	GCTTGGGTAGACACACTCTTGTGGTGTGTGTGTGGACGTAAACGCTGTGTGTGTG	174
QY	1741	ctgcttggagactgtcgtcgtgttctgccttgacactagacacccctgtgtgtgaatgcaga	180
Db	1741	CTGCTTGGGACGTGTGTGCTGCTGTTGTGCTGTGCACCTTAAGAACCCGTGTGTGAGTCACA	180
QY	1801	ggggggccgcctgtccttctctatgcctgggggtccctctggcaagagatgtgtgagccctcat	186
Db	1801	GGGGGGCCGCTGTCTTCTTATGTGTGTGGCTCCCTGTGGACGAGATGTGTGACGCTCAT	186
QY	1861	ggcttcttggggaaaccacaagcctgcgtgtgtcctacgcgcagccctcttgcctt	192
Db	1861	GGCTTCTTTTGGGAAACCAAGGCGCTGCGCTGTCTACGCCCAAGGCGCTCTTGTGCCCTT	192
QY	1921	ggttcaacactcttcctgtcctcgtcgcgaaggttgtaattccaactaatatcatcttc	198
Db	1921	GGTTCAACACTTTCCTGTCTCCTGTGACAGTGTGCTCAATTCCAACTATATCATCTTC	198
QY	1981	aagtttccacaaggaatccatacatctcaacaagccttggtgttcaaaaacaacgctgtgc	204
Db	1981	AAGTTTCCACAAAGGATCTACATTCTACACAGCCTGGGTCCAAAAACACAGGTGTGTGC	204
QY	2041	ctgttgtgatgatcagctcagcgcgcacagctgtctatctgtctcaacttggcctgtgtgtg	210
Db	2041	CTGTGTGTGATGATCAGCTCAGCGGCCACAGCTCTATCTGTCTAATCTGGCTGTGTGTG	210

QY	2101	tgagccaccagctgcgtctagggaataataccaagccttcccccacatcgtgtatgctttagtgc	2160
Db	2101	ttgaccctccacctccctgctaagggaataccaagccttcccccacatcgtgtatgctttagtgc	2160
QY	2161	acaagagaccaactccctcgtggtctcactatcagccttcctccatcagatgacctccctccatc	2220
Db	2161	ACAAGAGACCAACTCCCTGGGCTTCATACAGGCGCTTCCTACATAGGCTCCTCTTCATAC	2220
QY	2221	atgctccttgcctcgtcagctactccctggtgtaagaaacttgcacagaagaactacaacagagccaaa	2280
Db	2221	ACTGCGCTTTGGCTGCGAGCTACCTGGGTAAAGACTTGCCAGAGAACTACAAACAGGGCANA	2280
QY	2281	tggtgtaaccttaagcctcgtctcttaactctgtgtcctgtatgctgccttcttcaaccagcgc	2340
Db	2281	TGtGTACACCTTCAGGCTGCTCTTCACAACTTCGTCTCTGATGCGCTTCCTTCACCAAGGGCC	2340
QY	2341	agcgcctacagcagcgaagtaactcgtcgtcgcgcacaataatgactcgtgaggtatgagagacctg	2400
Db	2341	AGCGCTACAGACGGAGATACCTGCTCGGGCCAAATATGTGGCTTGAGCAACCTTG	2400
QY	2401	agcaagcgccttcgtcgtggtatcttctcctaagtgactagtgatcctcctcgcgcaccaagac	2460
Db	2401	AGCAGCGGGTTCGGTGGGATTTTCTGCTAGTGTGATGATCTCTGCGGCCACAGAC	2460
QY	2461	ctcaacagcacagagcaactccagcgcctccatcagagactacagagcgcgtgcgtctcc	2520
Db	2461	CTCAACAGCACAGACACTTCCAGGCGCTCCATTCACAGACTACACGAGCGCGCTCGCGCTCC	2520
QY	2521	acctga 2526	
Db	2521	ACCTGA 2526	
RESULT	2		
LOCUS	AX823900	2951 bp	DNA linear PAT 02-NOV-2001
DEFINITION	Sequence 9 from Patent WO0172842.		
ACCESSION	AX823900		
VERSION	AX823900.1	GI:16609872	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	human.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Maize, D.W., Scoville, J., Donoho, G., Turner, C.A., Friedrich, G.C.,		
JOURNAL	Abidin, A., Zambrowicz, B. and Sands, A.T.		
FEATURES	Novel human 7tm proteins and polynucleotides encoding the same		
source	Patent: WO 0172842-A 9 04-OCT-2001;		
	Lexicon Genetics Incorporated (US)		
	location/Qualifiers		
	1..2951		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	559 a 888 c 831 g 669 t 4 others		
ORIGIN			
Query Match	100.0%; Score 2526; DB 6; Length 2951;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 2526; Conservative	0; Mismatches	0; Indels	0; Gaps
QY	1	atgcgcctcgtcagcgcctgcctgctggtgcgcgcgcagctctcatctcgcgtcgtgcgc	60
Db	232	ATGCGCTTCGACAGGCTGCTGCTGGCGGCCCGACGCTTCATTTCCGCGCTGCGGCC	291
QY	61	tttgcctgcataagcaagagatctctcctcgaactcaaccctcccgagatatactctct	120
Db	292	TTTGGCTGCAATAGACAGGAGTCTTCCTCTGACTTCACCCCTCCCGAGATTAACCTCG	351
QY	121	gaagcgcgttccctcctcattctgctgctgtcgcagaggtgagacagagccagtgtagc	180
Db	352	GCAGGCGTGTTCCTCCATTTCTGAGTGTCTGCAAGGTGAGGCACAGACCTCAGGTGAGCC	411

QY	181	ctgtgtgacaggtcttgttagcttcaatagcatgtgtctacacccctcttcacgagctacgtg	240
Db	412	CTGTGTGACAGGCTTGTACTCTTCATATGACATGACATGCGCTTACACCTCTTCCAGGCTATTCGGG	471
QY	241	cttggaggtttggagagataaacaactccacagccctctctccaaatcacaccccttgggtac	300
Db	472	CTTGGGGTTGGAGATTAACAACTCCACG66CCTCTCTCCAAATCACCTTGGGGTAC	531
QY	301	cagctgtatgtatgtgttcttgcactctgcacaatgttatgtcacagcttagagtgtctcc	360
Db	532	CAGCTGTATGATGTGTGTATTGTAGACTGTGGCAATGTGTATGCAACGCTTAGAGTGTCTCC	591
QY	361	ctgcacagggacaacacacatagagcttccaaaggagacctctccactatccctcaggtg	420
Db	592	CTGCACAGGGACAACACCATAGAGCTTCCAGGAGACCTTCTCCATATTCCTCTACGGTG	651
QY	421	ctggagatgtatggccttgtagcaagcacacacgtgtctgcacacaaagccgctgtctagc	480
Db	652	CTGGAGATGTATGGGCTCTGACAGACCAACCGTGTCCACACACAGCGCCTGTGTAGC	711
QY	481	cccttcctcgtgtgcccatgatatagctatcgccagcagcagagacgtcacgcttgaaggg	540
Db	712	CCTTCTCCGTGGTCCCATGATTACTATGCGGCGACGAGGAAACGCTCAGCGTAAAGGG	771
QY	541	cagtatccctctctctctgcgcacacatccccaatgacaagtaccaggtgagaccatgtg	600
Db	772	CAGTATCCCTCTTCTCTGCGACACATCCCAATGACAACTACAGATGAGACCATGATGGT	831
QY	601	ctgtctgtctgcagaagttccgggtgagacctgtgattctctctgtgttgacgaatgtgacgactat	660
Db	832	CTGTCTGTCTGCAGAAAGTTCCGGGTGAGCTGTGATCTCTGTGGTGACGATGCACACTAT	891
QY	661	gggcagctcaaggggtgcagagcactgtgagaaacagcgcacgtcgaaggatctgcatgct	720
Db	892	GGGCAAGTAAAGGGGTGCAGAGCATGTGAGAAACAGGCCACTGTCTCAGGGATCTGCATTGCT	951
QY	721	ttcaagagacatcatgacctctctctccacagggtgggcgatgtagagagatgcaatgctcatg	780
Db	952	TTCAAGGACATCATGACCTTCTCTCGCCACAGTGGGCGATGAGAGATGCAGTGCCTCATG	1011
QY	781	cgccacactggcccacagcggcgggcgaacocgtcgtgtgttttccagccgggaattggcc	840
Db	1012	CGCCACTCGGCCCAAGGGCCGGGGCACCGTGTGTGTTTTCAGCGCGGCAATGTGGCC	1071
QY	841	aggggtgttttccgagtcctcgtgtgcctgcacacacctaactgcgcgaagtgttggtctcccca	900
Db	1072	AGGGGTGTTTTCAGTGTCCGTGTGCTGTGACCAACTACTGCGAAGTGTGGGTGCCCTCA	1131
QY	901	gaagccttggccctctccagcgcacatcaactgtgggtggtccgggataccagcgcatttggatg	960
Db	1132	GAAAGCTTGGGCCCTTCTCCAGGACAACTACATCGGGGTGCCGGGATCCAGGCGCATTTGGATG	1191
QY	961	gtgtcttggcgttggccaatccagaagagggtgttccctcgtgcctgaagcgttttgaagaagcc	1020
Db	1192	GTGTCTTGGCGCTTGGCCATCCAGAAAGAGGCGCTGTCCGTCCATMAAGCGTTTGAATAAACCC	1251
QY	1021	tatgcccgggagacagaagggggcccttaaggccttgcacaaagggtctccgtgtggagcagc	1080
Db	1252	TATGCCCGGCGACGACAGAGAGGCCCTTAGGCTTGGCCACAAAGGGCTCTGCTGTGCACGACC	1311
QY	1081	aatcagctcttgcagaagaatgcacaagcttllcatgttccacacagatgccaaagctcacaagcc	1140
Db	1312	AATCAGCTCTGCAGAGAAGATGCAAGCGTTTCATGRCACACAGATGCCCAAGTCAAAACC	1371
QY	1141	ttctccatgagttcttgccttacaagcatcacgggtctgtatgtggttggccatgtgctc	1200
Db	1372	TTTCTCCATGAGATTTCTGCTTCAACAGCATACCGGGCTGTGTATGCGGTGGCCATCGGCTCTC	1431
QY	1201	cacacagctccttggggtcttgcctcttggagcttgttttccaaaggccgggtcttaacccctggag	1260
Db	1432	CACCAAGCTCTTGGGCTTGGCTCTGGAGGCTGTGTTCCACAGGGCGGAGTCTACCCCTGGGAG	1491

1261 cttttgagcaatccacaagatgcatcttctctacacagaacatgctgagctttaa 1320
 1492 cttttgagcaatccacaagatgcatcttctctacacagaacatgctgagctttaa 1551
 1321 gacaaagagatccctctagtagctatacaaatgtcctggaatggaatggaatgga 1380
 1552 gacaaagagatccctctagtagctatacaaatgtcctggaatggaatggaatgga 1611
 1381 tggaccttaccggtctctggtctctacatgctctcagttcagttcagttcagttcag 1440
 1612 tggaccttaccggtctctggtctctacatgctctcagttcagttcagttcagttcag 1671
 1441 accaaatccagtgagcaaggaagaaacagatgctcctcagttcagttcagttcag 1500
 1672 accaaatccagtgagcaaggaagaaacagatgctcctcagttcagttcagttcag 1731
 1501 tctcttgaagggcaccaggaatggttgaagtttccatcactgctgcttgaagttgtg 1560
 1732 tctcttgaagggcaccaggaatggttgaagtttccatcactgctgcttgaagttgtg 1791
 1561 cctctgagggctgagacatctcctcaacaagatgacatcagatgagcagccttgag 1620
 1792 cctctgagggctgagacatctcctcaacaagatgacatcagatgagcagccttgag 1851
 1621 aaagaagatgagcagcctgaggaagacagacatctcctcagcagcctgagttgtt 1680
 1852 aaagaagatgagcagcctgaggaagacagacatctcctcagcagcctgagttgtt 1911
 1681 gcttggcgtgagcaac 1740
 1912 gcttggcgtgagcaac 1971
 1741 ctgcttggagcctgagcctgagcctgagcctgagcctgagcctgagcctgagcctg 1800
 1972 ctgcttggagcctgagcctgagcctgagcctgagcctgagcctgagcctgagcctg 2031
 1801 gggggcggcctgagcctgagcctgagcctgagcctgagcctgagcctgagcctg 1860
 2032 gggggcggcctgagcctgagcctgagcctgagcctgagcctgagcctgagcctg 2091
 1861 gggcttcttggaggaacccaagggcctgagcctgagcctgagcctgagcctgagcct 1920
 2092 gggcttcttggaggaacccaagggcctgagcctgagcctgagcctgagcctgagcct 2151
 1921 ggttccacatctcctgctgctgctgagcctgagcctgagcctgagcctgagcctg 1980
 2152 ggttccacatctcctgctgctgctgagcctgagcctgagcctgagcctgagcctg 2211
 1981 aagtttccacaaggaatcactatctcacaagcctgagcctgagcctgagcctgagc 2040
 2212 aagtttccacaaggaatcactatctcacaagcctgagcctgagcctgagcctgagc 2271
 2041 ctgcttggatgatcagcctgagcctgagcctgagcctgagcctgagcctgagcctg 2100
 2272 ctgcttggatgatcagcctgagcctgagcctgagcctgagcctgagcctgagcctg 2331
 2101 tggacccaatgctgctgagcctgagcctgagcctgagcctgagcctgagcctgagc 2160
 2332 tggacccaatgctgctgagcctgagcctgagcctgagcctgagcctgagcctgagc 2391
 2161 aagaaggaacaaatcctgagcctgagcctgagcctgagcctgagcctgagcctgagc 2220
 2392 aagaaggaacaaatcctgagcctgagcctgagcctgagcctgagcctgagcctgagc 2451
 2221 agtgccctgagcctgagcctgagcctgagcctgagcctgagcctgagcctgagcct 2280
 2452 agtgccctgagcctgagcctgagcctgagcctgagcctgagcctgagcctgagcct 2511
 2281 tggatccatcagcctgagcctgagcctgagcctgagcctgagcctgagcctgagcct 2340
 2512 tggatccatcagcctgagcctgagcctgagcctgagcctgagcctgagcctgagcct 2571
 2341 agcgttcaagcagcaggaatcagcctgagcctgagcctgagcctgagcctgagcct 2400

Db 2572 AGGCTTACAGACGGCAAGTACCTGCTGGGCCCAACATGATGCTGGGCTGACACCTG 2631
 Qy 2401 agcagcgtctgagtgatatttctgctaaagtatgtagtctctgcccagc 2460
 Db 2632 AGCAGCGCTCGGTGGATTTTCTGCTTAAGTGTACGTGATCTGCTGCCCCAGAC 2691
 Qy 2461 ctcaacagcaacagacatcagcctcctcattcaagactacagagcgtgagctcc 2520
 Db 2692 CTCACAGACAGACGACTTCAGGCGCTTCATTCAGAGACTACAGAGGCGCTCGGCTCC 2751
 Qy 2521 acctga 2526
 Db 2752 ACCTGA 2757
 RESULT 3
 AX354027 2759 bp DNA linear PAT 06-FEB-2002
 LOCUS Sequence 13 from Patent WO0203848.
 DEFINITION AX354027
 ACCESSION AX354027
 VERSION AX354027.1 GI:18618964
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Naito A.T.
 TITLE Method for opening the blood-brain barrier
 JOURNAL Patent: WO 0203848-A 13 17-JAN-2002;
 NAITO, Albert T. (US)
 FEATURES
 source Location/Qualifiers
 1..2759
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID NO: 7475203CB1"
 BASE COUNT 523 a 828 c 782 g 626 t
 ORIGIN
 Query Match 91.9%; Score 2322; DB 6; Length 2759;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2522; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 atgctgctcagcagcctgctgctgctgctgctgctgctgctgctgctgctgctgctg 60
 Db 53 ATGCTGCTCGACAGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 112
 Qy 61 ttgctgcatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 120
 Db 113 TTGCTGCGCATGACAGGAGTCTTCTCGACTTCAACCTCCCGAGATTAACCTCTG 172
 Qy 121 gcaagccttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 180
 Db 173 GCAGGCTGTGCTCTCTCCATTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 232
 Qy 181 ctgtgcaagcgtctgtgagctcctcctcctcctcctcctcctcctcctcctcctcctc 240
 Db 233 CTGTGTGACAGGTCTGTGAGCTTCAATGAGCATGAGCTTACACCTCTTCCAGGCTATGCG 292
 Qy 241 ctgtggttggaggaataaacaactcagcagcctcctcctcctcctcctcctcctcctc 300
 Db 293 CTGCGGCTTGAAGAGATAAACATCTCAGCGGCTCTGCTGCCAATCACCCCTGGGGTAC 352
 Qy 301 cagcttatagatggt 360
 Db 353 CAGCTGTATGATGT 412
 Qy 361 ctgcagagcaaac 420
 Db 413 CTGCCAGGGCAACCAACATAGAGCTCAAGAGAGACCTTCTCAGCTATTCCTTACGAGTG 472

421 ctgagcagtgatggcgctgacagcaaccaacgctgtctccaccacagcgccctgtcgaac 480
|||||
Db 473 CTGGAGATGATGGGGCTGACAGCAACCGTGTGCTGCACACAGCCCGCTGTGAGC 532
OY ccttcctctgtgtcccatgattagctatgctgagcagcagcagcagcgtcagcgttgaacg 540
|||||
Db 533 CTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 592
OY 541 cagttccct 600
|||||
Db 593 CAGTATCCCTCTTCTCTGCGACACATCCCAATGACAAATACAGTGGAGACCAATGAGT 652
OY ctgctctctgacagaaagctcgtgtgagccttgatctctctgtgtgtgagcagctgacat 660
|||||
Db 653 CTGCTGCTGACAAATTTGGGTGACCTGATCTCTGTGTTGGACAGCTACGATCTAT 712
OY 661 gggcagcagtgaggtgacagcagctgagaaacagcgaactgtctcagtggaatctgtc 720
|||||
Db 713 GGGCAGCTAGGGGTGACAGGACCTGAGAAACAGGACAGCTGATGAGGGATCTGATGCT 772
OY 721 ttcaagagcatatgct 780
|||||
Db 773 TTCAAGGACATATCCCTTCTCTCTCCAGGTGGCGATGAGAGATGATGATGATGATGATG 832
OY 781 cgcacacctgagccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
|||||
Db 833 CGCCACCTGCGCCACAGCG 892
OY 841 aggtgtctctctcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 900
|||||
Db 893 AGGGGTGTTTTCGAGTCCGTGCTGCTGACCAACCTGACGACGACGACGACGACGACGAC 952
OY 901 gaagcctgagccct 960
|||||
Db 953 GAGGCTGAGCCCTCTTCAAGGACATCACTGGGGTGCCTGGGATCCAGGCTATGGGATG 1012
OY 961 gtctctgagcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020
|||||
Db 1013 GTGCTGGGGTGGCCATCCAGAAAGGGGTGTCCCTGGCGTGAAGCGCTTGAAGAGCG 1072
OY 1021 tatgcccgggagcaagaagagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
|||||
Db 1073 TATGCCCCGGGAGACAAAGGCCCTTAGGCCCTTCCACAAAGGGCTCCGTGTCAGCAG 1132
OY 1081 aatcagcctctgagagaatgcaagccttctcagcagcagcagcagcagcagcagcagc 1140
|||||
Db 1133 AATCAGCTCTGACAGAAATGCCAAGCTTTCATGACACACAGATGCCAAGCTCAAGAGC 1192
OY 1141 ttctccatgagttct 1200
|||||
Db 1193 TTCTCCATGAGTTTTCGCTACAAAGCATACCGGGCTGTGTATGGCGTGGCCATGGCCTC 1252
OY 1201 caacagcctctgagcgt 1260
|||||
Db 1253 CACAGCTCTCTGGGCTGTGCTCTGAGCTTCTCCAGGGCGCGAGTCAACCCCTGGCAG 1312
OY 1261 cttttggagcagaatcacaaggtgattctctctctctctctctctctctctctctct 1320
|||||
Db 1313 CTTTGGAGCAGATCCACAGGATGCTTCTTCAACAGAGACCTGTGGCGTTTAAAT 1372
OY 1321 gacaaacagatccctctcagtagctataataatctgctctgagcagctgagcagcagc 1380
|||||
Db 1373 GACAAACAGATCCCTCAGTAGTATATAATTTGCTGGAGCTGGAATGAGACCCAG 1432
OY 1381 tggacctcagtgct 1440
|||||
Db 1433 TGGACCTTCAAGCTCTGTGTTCTTCAACATGCTCCAGTTCACACTAAATATGAG 1492
OY 1441 accaaatccagtgacaggaagagacaaacagcagcagcagcagcagcagcagcagc 1500
|||||
Db 1493 ACCAAATTCAGTGGCAGAGAAAGACACAGGTGCTTAAAGTGTGTGTCAGCGAGC 1552
OY 1501 tgtctgaagggcaccagcagtggttaccggtttccatcactgctgttgagtggtg 1560

|||||
Db 1553 TGTCTTGAAGGGGACACAGGAGTGTAGGGTTTCATCACTGCTGTAGTGTG 1612
OY 1561 cccctgtgagcgttgagccttccctcaacagatgacctctacagatgacagcctgtgag 1620
|||||
Db 1613 CCTGTGGGGCTGGGACCTTCTCAACAGAGTGAACCTTACAGATGACAGCCTTGTGG 1672
OY 1621 aaagaagatgggaccttgaaggaaagcagcagcagcagcagcagcagcagcagcagc 1680
|||||
Db 1673 AAGAGAGTGGGACCTAGGGAAGCCAGACCTGCTTCCGCGACCTGTGTGTTTG 1732
OY 1681 gcttctgagcagacacccctctctctctctctctctctctctctctctctctctct 1740
|||||
Db 1733 GCTTGGCGTACACACCTCTTGGGTGCTGCGGACGATACAGCGTGTGCTGTGCTG 1792
OY 1741 ctgcttgagcagcgtgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1800
|||||
Db 1793 CTGCTGGGAGTGTGGGCTGCTTGTGCTGACCTGACACCCCTGTGTGAGTACGA 1852
OY 1801 gggggccgctgtgtct 1860
|||||
Db 1853 GGGGGCCGCTGTGCTTGTATGCTGGGCTCCCTGACAGATGATGACGCTCTAT 1912
OY 1861 ggtctcttggggaaccccaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1920
|||||
Db 1913 GGTCTCTTGGGGAACCCACAAAGCGCTGGCTGTGCTACAGCAGGCGCTTGTGCGCTT 1972
OY 1921 ggttcaacatct 1980
|||||
Db 1973 GGTTCACCATCTTCTGCTGCTGACAGTGTGCTATTCACATATCATCATCTTC 2032
OY 1981 aagtttccacaaaggtactactactactactactactactactactactactactact 2040
|||||
Db 2033 AAGTTTCCACAAAGTACCTACATTTCAACGCTGTGGTCCAAACACAGGCTGTGGC 2092
OY 2041 ctgttctgagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2100
|||||
Db 2093 CTGTTGTATGATACGCTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2152
OY 2101 tggaccccaactctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2160
|||||
Db 2153 TGGACCCCACTCTGCTGAGGAATACAGCGCTTCCCATCTGTGTGATGCTGAGTGC 2212
OY 2161 acagaagacaaactccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2220
|||||
Db 2213 ACAGAGACCAACTCCCTGGGCTTCACTAGGCTTCTCTACAAATGCGCTCTCTCATC 2272
OY 2221 agtgcccttgcctgacagcagcagcagcagcagcagcagcagcagcagcagcagc 2280
|||||
Db 2273 AGTGCTTGTGCTGACAGTACCTGGGTAGGACTTGCAGAGAACTAACAGAGGCCA 2332
OY 2281 tgytcaaccttaagcctgtctctcaactctgtgtgtgtgtgtgtgtgtgtgtgtgt 2340
|||||
Db 2333 TGTGTACCTTCAAGCTGCTTCAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2392
OY 2341 agcgtctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2400
|||||
Db 2393 AGCGTCTACAGAGGCAAGTACCTGCTGCGGCGCACATATGCTGTGAGACAGCTG 2452
OY 2401 agcagcagcgtctgt 2460
|||||
Db 2453 AGCAGCGCTTGGGTGATTTTCTAGTATGCTAGTATGCTGTGCGGCCAGAC 2512
OY 2461 ctcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2520
|||||
Db 2513 CTCAACAGCAGAGACTTCCAGGCGCTCATTCAGGACTACAGAGGCGGTGCGGCTCC 2572
OY 2521 acctga 2526
|||||
Db 2573 ACCTGA 2578

RESULT 4

Db 1741 ATCTTCAAGTTTTCACCAAGGTACTACATTTCACGAGCCTGGGTCCAAACACAGGT 1800
Qy 2035 gctgacctgtgtgtgatgatcagctcagcgccagcgtctatcttctactgtgtg 2094
Db 1801 GCTGGCCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
Qy 2095 gtgtgtgtgagaccac 2154
Db 1861 GTGTGTGTGAGACCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Qy 2155 gattgacacagac 2214
Db 1921 GAGGTGACAGAGACCAACTCCCTGGGCTTCATGCTGCTTCCTGCTGCTGCTGCT 1980
Qy 2215 tccatcagctgt 2274
Db 1981 TCCATGATGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Qy 2275 gcccaatgtgtac 2334
Db 2041 GCCAAATGTGTACCTTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Qy 2335 accgac 2394
Db 2101 ACCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Qy 2395 agcctgtgac 2454
Db 2161 AGCCTGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Qy 2455 ccagac 2514
Db 2221 CCAGACCTCAAC 2280
Qy 2515 ggcctccac 2574
Db 2281 GGTCTCACCTGTA 2292

RESULT 5
AX282896
LOCUS AX282896 1101 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 5 from Patent W00172842.
ACCESSION AX282896
VERSION AX282896.1 GI:16609870
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
REFERENCE
AUTHORS Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,
Abidin,A., Zambrowicz,B. and Sands,A.T.
TITLE Novel human 7tm proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0172842-A 5 04-OCR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1. 1101
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 219 a 308 c 323 g 250 t 1 others
ORIGIN

Query Match 39.5%; Score 999; DB 6; Length 1101;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 595 atgt 654
Db 1 ATGT 60
Qy 655 gactatggcagctagggtgtcagcgacacacacacacacacacacacacacacacacacac 714

Db 61 GACTATGGGACGTAGGGGTGACAGGACATGGAGAAACAGGGCCAGCTGCTCAGGGATCTGC 120
Qy 715 atgtcttcaagac 774
Db 121 ATTGCTTCAAGACATCATCT 180
Qy 775 ctcatgtgcaac 834
Db 181 CTGATGCGCCACCTGCGCCAGGCGGGCCACCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Qy 835 ttgtgcaaggt 894
Db 241 TTGGGCAAGGT 300
Qy 895 gctcagaagcctgt 954
Db 301 GCTTCAGAAAGCCTGGGCTCTCCAGGACATCACTGTGGGTGCGCGGATCCAGGCAATT 360
Qy 955 ggaattgt 1014
Db 361 GGGATGCTGT 420
Qy 1015 gaagcctatgt 1074
Db 421 GAAGCCTATGCGCGGAGACAGAGAGGCGCTTGTGCTTCCACAAAGGCTCTGCTGTGC 480
Qy 1075 agcgaac 1134
Db 481 AGCAGCAATCAGCTCTGTGAGAGAAATGCCAACCTTCAATGACACACATCCCAAGCTC 540
Qy 1135 aaagccttccac 1194
Db 541 AAAGCCTTCTCATGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy 1195 ggcctccac 1254
Db 601 GGCTCCACACAGCTCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Qy 1255 tggcagcttttggagcagac 1314
Db 661 TGGAGCTTTTGGAGCAATCCACAAAGGTGATTTCTTCAACAAAGCACTGTGTGC 720
Qy 1315 ttatgac 1374
Db 721 TTTATGACAAACAGATCCCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 780
Qy 1375 cccaagtgac 1434
Db 781 CCCAAGTGAGACTTCAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy 1435 aatgagac 1494
Db 841 AATGAGACCAAAATCCAGT 900
Qy 1495 agcagcgtcttgaagggac 1554
Db 901 AGCAGCTGTCTTGAAGGACACAGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Qy 1555 ttgtgcccgt 1593
Db 961 TGTGTGCTGT 999

RESULT 6
AL591866
LOCUS AL591866 102539 bp DNA linear PRI 25-OCR-2001
DEFINITION Human DNA sequence from clone RP11-58A11 on chromosome 1, complete
sequence.
ACCESSION AL591866
VERSION AL591866.13 GI:16501194
KEYWORDS HTG.
SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 102539)
 AUTHORS Wallis, J.
 TITLE Direct Submission
 JOURNAL Submitted (25-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk
 On Oct 26, 2001 this sequence version replaced gi:15022299.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; Sw:
 SWISSPROT; Tr: TREMBL; Wp: WormPEP; Information on the WormPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr1
 RP11-58A11 is from the library RPEC1-11.1 constructed by the group
 of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-58A11. It may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true right end of clone RP11-58A11 is at 102539 in this
 sequence. The true left end of clone RP11-239P22 is at 13247 in
 this sequence. The true right end of clone RP4-650H14 is at 2000 in
 this sequence.

FEATURES
 source
 1..102539
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-58A11"
 /clone_1lb="RPEC1-11.1"
 582..618
 /note="Sequence from overlapping clone RP4-650H14
 (AL158217). Assembly confirmed by restriction digest."
 1452..1473
 /note="Sequence from overlapping clone RP4-650H14
 (AL158217). Assembly confirmed by restriction digest."
 48541..48603
 /note="Sequence from overlapping clone RP11-239P22
 (AL159177). Assembly confirmed by restriction digest."
 69007..69267
 /note="Sequence from overlapping clone RP11-239P22
 (AL159177). Assembly confirmed by restriction digest."
 100154..100179
 /note="Sequence from overlapping clone RP11-239P22
 (AL159177). Assembly confirmed by restriction digest."
 BASE COUNT 25289 a 26768 c 27360 g 23122 t
 ORIGIN
 Query Match 34.9%; Score 882; DB 9; Length 102539;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 932; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1594 gacctacagatgccaagctctgttggaagaagatggaacactgaggaagccagacc 1653
 |||||||
 Db 75842 GACCTCTACAGATGCCAGCTTGTTGGGAAGAAGACTGGGACACTGAGGAAGCCAGACC 75901
 QY 1654 tgctccgcgcgaactgtgtgtttttgtgcttgcgtgaacacacactcttggtgtctgcg 1713
 |||||||
 Db 75902 TGCTTCCCGCAGCACTGTGTGTTTTGGGTTTGGCGGACACACCTCTTGGGTGCTGCG 75961
 QY 1714 ggaagtaaacgctgtgtctgtctgtctgtctgtggaacgcctgagcctgttgcctgac 1773
 |||||||
 Db 75962 GCAGCTAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 76021
 QY 1774 ctgaagaccctgtgtgtgaagtaagcaagggcgccgtgtgtcttctatgtctggtctc 1833
 |||||||
 Db 76022 CTGAGACACCCCTGTGTGTGAGGTGACGAGGGGCGCGCTGCTTGTATGCTGAGCTCC 76081
 QY 1834 ctggcgcgcggtgtgtgtgcaagcctctatgtctctcttggggaaacccaagcgctgtgc 1893
 |||||||
 Db 76082 CTGGCAGCAGGTAGTGGCAGCCTCTTATGCTTGTGGGGAACCCAGAGGCTGCGTGC 76141
 QY 1894 ttgctacgcgaagccctcttgcctgtgttcaaccatctcctgtcctgtacagt 1953
 |||||||
 Db 76142 TTGCTACGCGCAGGCCCTCTTGGCTTGGCTTGCATCTTCCGTGCTGCTGAGATT 76201
 QY 1954 cgtcatccaactaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatca 2013
 |||||||
 Db 76202 CGCTCATTCACACTAATCATCATCTTCAAGTTTCCACCAAGATACCTCATCTTCACTCCAC 76261
 QY 2014 gctgtgtccaagaacagaggt 2073
 |||||||
 Db 76262 GCTGTGTCCAAGAACAGAGGT 76321
 QY 2074 ctatctgtctaactgt 2133
 |||||||
 Db 76322 CTATCTGTCTAATCTTGT 76381
 QY 2134 ttccccaactgt 2193
 |||||||
 Db 76382 TTCCCCCACTGT 76441
 QY 2194 ttccctacaatgt 2253
 |||||||
 Db 76442 TTCTCTTACATGTGCT 76501
 QY 2254 ttgcacagaaactacaacagagcgaatgtgtacaccttaagcctgtcttcaactgtgt 2313
 |||||||
 Db 76502 TTGCCAGAGAACTACAACGAGGCGCAATGTGTACCTTCAAGCTTCAACTTCGTG 76561
 QY 2314 tctctgatgt 2373
 |||||||
 Db 76562 TTCTGTGATGCTCTTCTTCAACAGGCGCAGCTGTACGAGGCGCAAGTACCTGCTGCGGC 76621
 QY 2374 aacatgatgt 2433
 |||||||
 Db 76622 AACATGATGCTGTGCTGTGACAGACCTGTGACAGGCGCTTGGTGTGTGTGTGTGTGT 76681
 QY 2434 tgcctatgtatctctgtccgcgcgaacactcaacagacagagcaacttcaagcctcatt 2493
 |||||||
 Db 76682 TGCTAGGTGATCTTGTGCGGCCAGACCTCAACAGCAGACAGACACTTCACAGGCTCAATT 76741
 QY 2494 caggaactaacagagcgctgt 2526
 |||||||
 Db 76742 CAGGACTACAGGAGCGCTGCGGCTCCACCTGA 76774
 RESULT 7
 AC108123/c 191481 bp DNA linear HTG 25-JAN-2002
 LOCUS
 DEFINITION Homo sapiens chromosome 16 clone RP11-56511, WORKING DRAFT
 SEQUENCE, 10 unordered pieces.
 AC108123
 AC108123.1 GI:18369969
 HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 191481)
TITLE DOE Joint Genome Institute.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 191481)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (25-JAN-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 623518
Center clone name: RPCI-11_56511

Summary Statistics
Consensus quality: 182477 bases at least Q40
Consensus quality: 187876 bases at least Q30
Consensus quality: 188750 bases at least Q20
Estimated insert size: 192000; agarose-fp estimation
Estimated insert size: 190581; sum-of-contigs estimation
Quality coverage: 6.71 in Q20 bases; agarose-fp estimation
Quality coverage: 6.76 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1230: contig of 1230 bp in length
* 1231 1330: gap of unknown length
* 1331 3606: contig of 2276 bp in length
* 3607 3706: gap of unknown length
* 3707 6459: contig of 2753 bp in length
* 6460 6559: gap of unknown length
* 6560 10289: contig of 3730 bp in length
* 10280 10389: gap of unknown length
* 10390 16864: contig of 6475 bp in length
* 16865 16964: gap of unknown length
* 16965 27807: contig of 10843 bp in length
* 27808 27907: gap of unknown length
* 27908 39098: contig of 11191 bp in length
* 39099 39198: gap of unknown length
* 39199 76534: contig of 37336 bp in length
* 76535 76634: gap of unknown length
* 76635 121347: contig of 44713 bp in length
* 121348 121447: gap of unknown length
* 121448 191481: contig of 70034 bp in length.
Location/Qualifiers
1. 191481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-56511"
/clone_lib="RPCI human BAC library 11"
BASE COUNT 48097 a 47075 c 46675 g 48729 t 901 others
ORIGIN
Query Match 34.9% Score 882; DB 2; Length 191481;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 932; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1594 gacctctacagatgccagcctgtgtggaagaagagtgagcactgaaggaagcagacc 1653

Db 72598 GACCTCTACAGATGCCACCTTTGTGGGAAGAAGAGTGGGACCGAGGGAAGCAGAC 72539
Qy 1654 tgcctccgagcactgtgtgttcttggcttctgtgtgagcaacactctgtgtctgt 1713
Db 72538 TGCTTCCCGCCAGCTGT 72479
Qy 1714 gcaactaacacgcctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 1773
Db 72478 GCAGCTAACACGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 72419
Qy 1774 cttagaacccctgt 1833
Db 72418 CTAGACACCCCTGT 72359
Qy 1834 ctgtgacagagtgatgtgagcactctatgtctcttctgtgtgtgtgtgtgtgtgt 1893
Db 72358 CTGGCAGCAGGTAGTGGGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 72299
Qy 1894 ttgtacgagcagggccctcttggccctgtgttcaacatctctctgtctgtctgtgt 1953
Db 72298 TTGCTACGCCAGGCCCTTTTGGCCCTTGTTCACACATCTTCTGTCTGTCTGTCTGT 72239
Qy 1954 cgtcatctcaactaatcatcatcttcaagtcttccacaaagtactatcttaccac 2013
Db 72238 CGCTCATTCCAACTAATCATCATCATCTTCAAGTTTCCACCAAGTACATCAATTCACAC 72179
Qy 2014 gctgtgtgtcaaaacacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2073
Db 72178 GCTGTGTGTCAAAACACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 72119
Qy 2074 ctctctctcaactgt 2133
Db 72118 CTATCTCTCACTGT 72059
Qy 2134 ttcccccactgt 2193
Db 72058 TTCCCCCATCTGT 71999
Qy 2194 ttctctcaaatgt 2253
Db 71998 TTCTCTCTCAATGT 71939
Qy 2254 ttgcagaggaactcaaaagaggaaggaaggaaggaaggaaggaaggaaggaagga 2313
Db 71938 TTGCCAGAGAACTACAAAGAGGCCAAATGTGTCACTTCACTGTCTGTCTGTCTGTCT 71879
Qy 2314 ttcctgtatcgtctctcttcaacacagcgccagcgtctacagagcaagtacgtctgtgt 2373
Db 71878 TCTCGATCGGCTTCTTCTACACAGGCCGCGCTACAGAGGCAAGTACTGTCTGTGTGT 71819
Qy 2374 aacatgatgt 2433
Db 71818 AACATGATGAGGT 71759
Qy 2434 tgcctacgt 2493
Db 71758 TGCTACGTGATCTCTGT 71699
Qy 2494 cagagactaacagagcgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2526
Db 71698 CAGAGACTACAGAGGCGGT 71666
RESULT 8
AX282898 AX282898 705 bp DNA linear PAT 02-NOV-2001
LOCUS
DEFINITION Sequence 7 from Patent WO0172842.
ACCESSION AX282898
VERSION AX282898.1 GI:16609871
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE	Euryarchaea; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Homidae; Homo.				
AUTHORS	1 (sites)				
TITLE	Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C., Abidin,A., Zambrowicz,B. and Sands,A.T.				
JOURNAL	Novel human 7tm proteins and polynucleotides encoding the same Patent: WO 0172842-A 7 04-OCT-2001;				
FEATURES	Lexicon Genetics Incorporated (US)				
SOURCE	1. 705 Location/Qualifiers				
BASE COUNT	130 a 231 c 170 g 173 t 1 others				
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606"				
Query Match	27.9%; Score 705; DB 6; Length 705;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches	705; Conservative 0; Mismatches 0; Indels 0; Gaps 0				

LOCUS	DEFINITION	AC062024	185061 bp	DNA	linear	HTG 01-SEP-2000
OY	18222	atgcgggctcccttggaagcaagtagtggcaacccctatggtctcttcttgggaaccaca	1881			
Db	1	ATGCTGGGGCTCCCTTGACACAGTAGTAGGACCTCTAGAGCCTCTTTTGGGGAACCCACA	60			
OY	1882	aggcctcgctgtctgtaagcagcagccctcttggcccttggtttcaaccactctctgtcc	1941			
Db	61	AGGCTCGGTGCTTCTAGCGCCAGGCCCCCTTTGCCCTTGTTTTCACCATCTTCCTGTCC	120			
OY	1942	tggcggacaagttggtgtaattccaaataatcatcttcaagtttcccaagtaacct	2001			
Db	121	TGCGTGACAGTGGCTCATCTCCAACTCAATCATCTTCAATTTTCCACCAAGTACT	180			
OY	2002	acattctcaacacgcttgggttccaaaacacaggttgctgyccttlttltgtatgtacgtca	2061			
Db	181	ACATCTPACACAGCGCTGGGTCCAAAACACAGGTGCGGCTGTTTGATGTACAGCTCA	240			
OY	2062	gcggccacagctgtcttatctgttctaacttggctgtgtgtgtgttggaacccactgctctag	2121			
Db	241	GGGGCCAGCAGCTGCTTATCTGTCTAACTTGGCTGGTGGTGTGGACCCCACTGGCTGCTAGG	300			
OY	2122	gaatacgaagcgtctcccccactctgtgtatgcttggatgtgtgtgacagaagaacaaatccctgggc	2181			
Db	301	GAAATCCAGACGCTTCCCCCATCTGGTGAATGCTTGAAGTGCACAGAGACCAACTCCCTGGGC	360			
OY	2182	ttcatacttggccttccctctacaaatgagcctcctctccatcagtgaccttltgcttgcagctac	2241			
Db	361	TTTCATAGTGGCTCTTCCTCTACAAATGAGCCCTCTCCATCATGAGTGCTTTGCTGACACTAC	420			
OY	2242	ctgtggttaagagcttgcacagagaatacacaagaagccaatgylgtcaaccttcagcctgc	2301			
Db	421	CTGGGTAAAGGACTTCCACAGAGACTACAAACGAGGCCAAATGTGTCACTTCAGCGCTGC	480			
OY	2302	ttcaacttcgtctgcctgatgcgctcttccacaacagcgagcaggtctacagaggaagtac	2361			
Db	481	TTCAACTTCGTCTCTGTGATCGCTTCTTTCACCAAGGCCAGGCTCTACACAGGCAAGTAC	540			
OY	2362	ctgccttcgagcacaacatgatggtcttggagtcagcagccttgcagcagcgtcttgcgttggat	2421			
Db	541	CTGCCTGGGGCCAAACATGATGGCTGGGGTGACACAGCTGATGACGAGGGGCTTCGGTGGTAT	600			
OY	2422	ttcttgtcctaagtgctacgttgatctctcttgcgcgccaagacctcaacaagcagagaacttc	2481			
Db	601	TTTTCTGCTTAAGTGTACGTGATCTCTGCGCCGCCAGACCTCAACAGCACAAGCAGCACTTC	660			
OY	2482	caggcctcattcagactacacgagcgaggtctgcgctcacactga	2526			
Db	661	CAGGCTCATTCAGAGACTACACAGAGGGCTCGGCTCCACTGA	705			

ACCESSION	SEQUENCE, 30 unordered pieces.
VERSION	AC062024
KEYWORDS	AC062024.2 GI:9958200
SOURCE	HTG; HTGS_PHASE1; HTGS_DRAFT.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 185061)
AUTHORS	Waterston,R.H.
TITLE	The sequence of Homo sapiens clone
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 185061)
AUTHORS	Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (21-APR-2000) Genome Sequencing Center, Washington
COMMENT	MO 63108, USA on Sep 1, 2000 this sequence version replaced gi:7630906.

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH02339H20
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-Primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 168551 bases at least Q40
Consensus quality: 174217 bases at least Q30
Consensus quality: 177307 bases at least Q20
Insert size: 173000; agarose-fp
Insert size: 182161; sum-of-contigs
Quality coverage: 3.68 in Q20 bases; agarose-fp
Quality coverage: 3.68 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 30 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1990: contig of 1990 bp in length
* 1991 2090: gap of unknown length
* 2091 3876: contig of 1786 bp in length
* 3877 3976: gap of unknown length
* 3977 5964: contig of 1988 bp in length
* 5965 6065: gap of unknown length
* 6065 7612: contig of 1548 bp in length
* 7613 7712: gap of unknown length
* 7713 11594: contig of 3882 bp in length
* 11595 15489: contig of 3795 bp in length
* 15489 15589: gap of unknown length
* 15589 19085: contig of 3496 bp in length
* 19085 19185: gap of unknown length
* 19185 21398: contig of 2213 bp in length
* 21398 21498: gap of unknown length
* 21498 33593: contig of 12095 bp in length
* 33593 33693: gap of unknown length
* 33693 36499: contig of 2806 bp in length
* 36499 36500: gap of unknown length
* 36500 40210: contig of 3611 bp in length
* 40210 40310: gap of unknown length
* 40310 43308: contig of 2998 bp in length
* 43308 43408: gap of unknown length
* 43408 46757: contig of 3349 bp in length
* 46757 43409

```

*	46758	46857	gap of unknown length
*	46858	51133	contig of 4278 bp in length
*	51136	51235	gap of unknown length
*	51236	55293	contig of 4058 bp in length
*	55294	55293	gap of unknown length
*	55294	62010	contig of 6617 bp in length
*	62011	62110	gap of unknown length
*	62111	68999	contig of 6889 bp in length
*	69000	69099	gap of unknown length
*	69100	74515	contig of 5416 bp in length
*	74516	74615	gap of unknown length
*	74616	80987	contig of 6372 bp in length
*	80988	81087	gap of unknown length
*	81088	87983	contig of 6886 bp in length
*	87984	88083	gap of unknown length
*	88084	98428	contig of 10345 bp in length
*	98429	98528	gap of unknown length
*	98529	107911	contig of 9383 bp in length
*	107912	108011	gap of unknown length
*	108012	117076	contig of 9065 bp in length
*	117077	117176	gap of unknown length
*	117177	128937	contig of 11761 bp in length
*	128938	129037	gap of unknown length
*	129038	138066	contig of 9029 bp in length
*	138067	138166	gap of unknown length
*	138167	149333	contig of 11168 bp in length
*	149335	149434	gap of unknown length
*	149435	165184	contig of 15750 bp in length
*	165185	165284	gap of unknown length
*	165285	181704	contig of 16420 bp in length
*	181705	181804	gap of unknown length
*	181805	183334	contig of 1530 bp in length
*	183335	183335	gap of unknown length
*	183335	185061	contig of 1637 bp in length

```

misc_feature      /note="assembly_name:Contig25"
                  62111. .68999
/note="assembly_name:Contig26"
misc_feature      69100. .74515
/note="assembly_name:Contig27"
misc_feature      74616. .80987
/note="assembly_name:Contig28"
misc_feature      81088. .87983
/note="assembly_name:Contig29"
misc_feature      88084. .98428
/note="assembly_name:Contig30"
misc_feature      98529. .107911
/note="assembly_name:Contig31"
misc_feature      108012. .117076
/note="assembly_name:Contig32"
misc_feature      117177. .128937
/note="assembly_name:Contig33"
misc_feature      129038. .138066
/note="assembly_name:Contig34"
misc_feature      138167. .149334
/note="assembly_name:Contig35"
misc_feature      149435. .165184
/note="assembly_name:Contig37"
misc_feature      165285. .181704
/note="assembly_name:Contig38"
misc_feature      181805. .183334
/note="assembly_name:Contig8"
misc_feature      183435. .185061
/note="assembly_name:Contig9"
BASE COUNT      45659 a 44829 c 45758 g 45906 t 2909 others
ORIGIN

```

URES	Location/Qualifiers
source	1. 185061
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="1"
	/clone="RP11-239H20"
	1. 1990
misc_feature	/note="assembly_name:Contig11"
misc_feature	2091. 3876
	/note="assembly_name:Contig12"
misc_feature	3977. 5964
	/note="assembly_name:Contig13
misc_feature	6065. 7612
	/note="assembly_name:Contig14"
misc_feature	7713. 11594
	/note="assembly_name:Contig15"
misc_feature	11695. 15489
	/note="assembly_name:Contig16
misc_feature	15590. 19085
	/note="assembly_name:Contig17
misc_feature	19186. 21398
	/note="assembly_name:Contig18
	clone.end:sp6
	vector_side:left"
misc_feature	21499. 33593
	/note="assembly_name:Contig36
misc_feature	33694. 36499
	/note="assembly_name:Contig19
misc_feature	36600. 40210
	/note="assembly_name:Contig20
misc_feature	40311. 43308
	/note="assembly_name:Contig21"
misc_feature	43409. 46757
	/note="assembly_name:Contig22
misc_feature	46858. 51135
	/note="assembly_name:Contig23
misc_feature	51236. 55293
	/note="assembly_name:Contig24"
misc_feature	55354. 62010

Best Local Similarity 99.7%; Pred. No. 5e-300;			
Matches	719;	Conservative	0; Mismatches 1; Indels 1; Gaps 1;
QY	1806	ccgcctgtgtcttctatagtgtggtcccttgcagcaagtagtggcaagctctatgctt	1865
Db	33701	CCGCGTGGGCTTTTATATGCTGGGCTCCCTGGCAGAGTAGTAGGACCTTATGGCTT	33760
QY	1866	ctttggtggaaacccaagaagctcggtgtgtcttgaagccagagccctcttggccttgctt	1925
Db	33761	CTTTGGGGGAACCCACAAAGGCTCGCTGCTTGTATAGCCAGGCGCTTTTGGCCCTTGCTT	33820
QY	1926	gaccctttccctgtctctgtgcctgaagttgcgcattcccaacatcatcatcttaagt	1985
Db	33821	CACCACTTTCCTGCTCTGCTATCAGTTTGCTATTCACACTAATCATATCTTCAAGTT	33880
QY	1986	tcccaaccaagttacatacttctaacagcgctgggttccaaacaaccaagttgcgtgctgt	2045
Db	33881	TTCCACAAAGTACACTTCTTACACAGCGTGGGGCAAAAACAAGGTGCTGGCGCTGTT	33940
QY	2046	tgtgatgatcaagctaacggccagctgtctatctgtctacttgcgtggtgttgac	2105
Db	33941	TGTGATGATACACTCAGCGGCCACACTGCTTTCGTGCTTCACTTGTGGTGGTGTGGAC	34000
QY	2106	cccatctgcctgttaaggaaataccagcgcttcccacatctggtgatgtgaatgacaaa	2165
Db	34001	CCCACTGCGTGTAAAGGAATACAGAGGCTTCCCACTCTGGTGATGCTTGAATGCACAA	34060
QY	2166	gaccaactccctgggtctatatactgtgctcttctcttacaatgctccctcttcacatg	2225
Db	34061	GACCAACTCCCTGGGGCTTCATCTAGTGGCTTCCTCTAACAATGGCGCTTCCTCATCAGTGC	34120
QY	2226	ctttgcctgcagctaacctggtgaagacttgcagagaactacaacagagccaatggt	2285
Db	34121	CTTTGCTCTTCACACTACTCTGGGTAAAGACTTGCACAGAGAACTCAACAGAGCCAAATGTGT	34180
QY	2286	gacccttaagcctgctcttcaacttggtgtgcccgtatcgctctcttcacacggccaaggt	2345
Db	34181	CACCTTCAACCGCTTCTTCAACTTGTGTGCTCGAGATCGCTTCTTCCACACGGCAGCGT	34240

QY 2346 ctacagcgcaagctaccctcgtcgccgcaacatgatgctggctgagcaagcctgagcag 2405
 |||||||
 Db 34241 CTACGACGGCAAGTACTCTGCTGCGG-CAACATGATGGCTGGGCTGAGCAGCCTGAGCAG 34299
 QY 2406 cggcttcggtggtatcttcctgaagtcgtacgtacgtccctgcgcgcacactca 2465
 |||||||
 Db 34300 CGGCTTCGGTGGTATTTTCTGCTTAAGTACTGATCTCTGCGCCACACCTCA 34359
 QY 2466 cagcagagcacttcacagcctccatcaatcaagactacagagcgctcgccctccactg 2525
 |||||||
 Db 34360 CAGCAGACGAGCACTTCAGGCGCTCCATTCAGGACTACAGAGCGCTGGCGCTCACACTG 34419
 QY 2526 a 2526
 Db 34420 A 34420

RESULT 10
 LOCUS G09859 270 bp DNA linear STS 15-AUG-1995
 DEFINITION human STS CHLC.GCT15G02.P16729 clone GCT15G02, sequence tagged site.
 ACCESSION G09859
 VERSION G09859.1 GI:941708
 KEYWORDS STS; STS sequence; primer; sequence tagged site.
 SOURCE human vector-pUC19 host-E.coli dut+ung+ (DH10B) Marker Selected genomic DNA prepared from xy individual of French nationality.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 270)
 Murray,J., Sheffield,V., Weber,J.L., Duyk,G. and Buetow,K.H.
 Cooperative Human Linkage Center
 Unpublished
 Synonyms: GCT15G02, CHLC.GCT15G02.T16713
 Contact: Dr. Jeffrey C. Murray
 UofI
 The University of Iowa
 Department of Pediatrics, Iowa City, IA 52242, USA
 Tel: (319) 356-3508
 Fax: (319) 356-3347
 Email: jeff-murray@uiowa.edu

Primer A: CTCACACAGGGGTCTCTAG
 Primer B: TGGTGTTTTGGCTTTC
 STS size: 124
 PCR Profile:
 denature: 30 seconds at 94 degrees C
 annealing: 75 seconds at 55 degrees C
 extension: 15 seconds at 72 degrees C
 PCR cycles: 27
 extension: 6 minutes at 72 degrees C

Protocol:
 Template: 30ng genomic DNA
 Primer: each 1.5 pmole
 dNTPs: each 200 uM
 Taq Polymerase: 0.3 units
 Total Vol: 10 ul

Buffer:
 MgCl2: 1.5mM
 KCl: 50mM
 Tris: 10mM
 pH: 8.3
 Location/Qualifiers
 1..270
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

STS
 primer_bind 95..114
 primer_bind complement(201..218)
 BASE COUNT 66 a 90 c 75 g 39 t
 ORIGIN

Query Match 10.7%; Score 270; DB 11; Length 270;
 Best Local Similarity 100.0%; Pred. No. 3.3e-141;
 Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1618 gggaaagaagatgggacactgaggaagcagaactgtctccgcgacatgtgtgttt 1677
 |||||||
 Db 270 GGGAAAGAGAGAGGGGACCTGAGGAGACCAAGACCTGCTCCCGGCACTGTGGTGT 211
 QY 1678 ttggcttcgtgagacacactttgtgtcgtctgagacataaagcgtgtgtgtgt 1737
 |||||||
 Db 210 TTGGCTTGGTGAGCAGCACACCTCTGGGTGCTGCGACTACAGCGTGTCTGCTG 151
 QY 1738 ctgctgttggaactgtgtgctgtttgttgccttggaacaccccttggttgagtgca 1797
 |||||||
 Db 150 CTGCTCTTGGGAGCTGCTGCGCTGTGCTGCTGCGACCTAGACACCCCTGTGGTGA 91
 QY 1798 gcaaggagccgctgtgtcttcttatgtgtgtgtcccttgagcagagtgatgtgacgtc 1857
 |||||||
 Db 90 GCAGGGGGCGCGCTGCTTGTATGCTGCGCTCCCTGCGACAGTATGTCACACCTC 31

QY 1858 tatgtcttcttgggagaccacaagcct 1887
 |||||||
 Db 30 TATGCTCTTGGGAGACCCACAGGCT 1

RESULT 11
 LOCUS HSHFR3S01 3106 bp DNA linear PRI 27-SEP-1996
 DEFINITION Human Kruppel-related 3 (HKR3) gene, exons 1-3.
 ACCESSION U45324
 VERSION U45324.1 GI:1177226
 KEYWORDS
 SEGMENT
 SOURCE 1 of 2
 human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 3106)
 Maris,J.M., Jensen,S.J., Sulman,E.P., Beltinger,C.P., Gates,K., Allen,C., Biegel,J.A., Brodeur,G.M. and White,P.S.
 Cloning, chromosomal localization, physical mapping, and genomic characterization of HKR3
 Genomics 35 (2), 289-298 (1996)
 96299777

2 (bases 1 to 3106)
 Maris,J.M., Jensen,S.J., Sulman,E.P., Beltinger,C.P., Gates,K., Allen,C., Biegel,J.A., Brodeur,G.M. and White,P.S.
 Direct Submission
 Submitted (09-JAN-1996) John M. Maris, Oncology/Pediatrics, The Children's Hospital of Philadelphia, ARC 902D; 324 South 34th Street, Philadelphia, PA 19104-4318, USA

JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 AUTHORS
 TITLE
 JOURNAL

FEATURES
 source
 exon
 intron
 exon
 intron
 TATA_signal

Location/Qualifiers
 1..3106
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p36.3"
 /clone="P1-4184, P1-4185, P1-4186"
 685..770
 /gene="HKR3"
 /number=1
 771..1172
 /gene="HKR3"
 /number=1
 1173..1931
 /gene="HKR3"
 /number=2
 1932..2746
 /gene="HKR3"
 /number=2
 1992..1998


```
misc_feature /clone="RP11-239H20"
1..1990
/note="assembly_name:Contig11"
misc_feature 2091..3875
/note="assembly_name:Contig12"
misc_feature 3977..5964
/note="assembly_name:Contig13"
misc_feature 6065..7612
/note="assembly_name:Contig14"
misc_feature 7713..11594
/note="assembly_name:Contig15"
misc_feature 11695..115489
/note="assembly_name:Contig16"
misc_feature 15590..19085
/note="assembly_name:Contig17"
misc_feature 19186..21398
/note="assembly_name:Contig18"
misc_feature 21499..33593
vector_side:left"
misc_feature 33694..36499
/note="assembly_name:Contig36"
misc_feature 36600..40210
/note="assembly_name:Contig19"
misc_feature 40311..43308
/note="assembly_name:Contig20"
misc_feature 43409..46757
/note="assembly_name:Contig21"
misc_feature 46858..51135
/note="assembly_name:Contig22"
misc_feature 51236..55293
/note="assembly_name:Contig23"
misc_feature 55394..62010
/note="assembly_name:Contig24"
misc_feature 62111..68999
/note="assembly_name:Contig25"
misc_feature 69100..74515
/note="assembly_name:Contig26"
misc_feature 74616..80987
/note="assembly_name:Contig27"
misc_feature 81088..87983
/note="assembly_name:Contig28"
misc_feature 88084..98428
/note="assembly_name:Contig29"
misc_feature 98529..107911
/note="assembly_name:Contig30"
misc_feature 108012..117076
/note="assembly_name:Contig31"
misc_feature 117177..128937
/note="assembly_name:Contig32"
misc_feature 129038..138066
/note="assembly_name:Contig33"
misc_feature 138167..149334
/note="assembly_name:Contig34"
misc_feature 149435..165184
/note="assembly_name:Contig35"
misc_feature 165285..181704
/note="assembly_name:Contig37"
misc_feature 181805..183334
/note="assembly_name:Contig38"
misc_feature 183435..185061
/note="assembly_name:Contig8"
misc_feature /note="assembly_name:Contig9"
BASE COUNT 45659 a 44829 c 45758 g 45906 t 2909 others
ORIGIN
Query Match 6.8%; Score 172; DB 2; Length 185061;
Best Local Similarity 100.0%; Pred. No. 1.2e-85;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 37929 CTGCTGCGCCCTGACGCTTCATTTCTGCTGCTG66CCTTTGCCATACACGAG 37870
Qy 82 tcttcctgactaccctcccgagattacctctgagcgtcttcctccat 141
|||||
Db 37869 TCTTCTCCGACTTCACCTCCCGAGATTTACTCTCGCAGCGCTGTCTCCAT 37810
|||||
Qy 142 tctgctgtctgcaggtgagcagaccgaggtgacctgtgtacagt 193
|||||
Db 37809 TCTGGCTGTCTGACGTGACAGACACCGAGGTGACAGGT 37758
|||||

RESULT 13
AC108123 191481 bp DNA linear HTG 25-JAN-2002
LOCUS Homo sapiens chromosome 16 clone RP11-56511, WORKING DRAFT
DEFINITION SEQUENCE, 10 unordered pieces.
AC108123
AC108123.1 GI:18369969
VERSION AC108123.1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFTN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 191481)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 191481)
DOE Joint Genome Institute.
Direct Submission
Submitted (25-JAN-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 623518
Center clone name: RPCI-11_56511
-----
Summary Statistics
Consensus quality: 182477 bases at least Q40
Consensus quality: 187876 bases at least Q30
Consensus quality: 188750 bases at least Q20
Estimated insert size: 192000; agarose-fp estimation
Estimated insert size: 190581; sum-of-contigs estimation
Quality coverage: 6.71 in Q20 bases; agarose-fp estimation
Quality coverage: 6.76 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1230: contig of 1230 bp in length
* 1231 1330: gap of unknown length
* 1331 3606: contig of 2276 bp in length
* 3607 3706: gap of unknown length
* 3707 6459: contig of 2753 bp in length
* 6460 6559: gap of unknown length
* 6560 10289: contig of 3730 bp in length
* 10290 10389: gap of unknown length
* 10390 16864: contig of 6475 bp in length
* 16865 16965: gap of unknown length
* 16966 27807: contig of 10843 bp in length
* 27808 27907: gap of unknown length
* 27908 39098: contig of 11191 bp in length
* 39099 39198: gap of unknown length
* 39199 76534: contig of 37336 bp in length
* 76535 76634: gap of unknown length
```

```

VGSHAGCECEPCAGTFFELNNSELIHIOPOGTEBMAKESKSTCPPTVEFLAMEPIIS
LVLIAMTTLILLILVGTAGLPAHEHHPVVRACGRCLCFMLTSLVAGSCSPVSEFEE
PTVPACILRQPLRPSLGRALFELSCITTSFOLYIFRKSTKYPTPTRYMAONHAGLTV
IVSSTVHLICLIMLVMTPTREYORPPLVILLECTEVNVSGFLAFTNHLISIS
TFVCSYLGKELPENYNBAKCVTESLINFVSWIAFTPMASISYQSYLPAVNVLAGLTT
LSGFGSGYFLPKVCVILCRPELNTEHFQASIODYTRRCGTT"
BASE COUNT      479 a      749 c      676 g      616 t
ORIGIN

Query Match      1.3%; Score 32; DB 10; Length 2520;
Best Local Similarity 100.0%; Pred. No. 6.3e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2119 agggaaaccagcgcttccccaatcgtgat 2150
|||||
|||||
Db 2116 AGGGAATACACAGCGCTTCCCCATCTGGTGAT 2147

RESULT 15
AX251144/c      5491 bp      DNA      linear      PAT 05-OCT-2001
LOCUS
DEFINITION      Sequence 112 from Patent WO0168912.
ACCESSION      AX251144
VERSION
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct
artificial sequence.
1 (bases 1 to 5491)
Olek A., Piepenbrock C. and Berlin K.
Diagnosis of diseases associated with tumor suppressor genes and
oncogenes
Patent: WO 0168912-A 112 20-SEP-2001;
EpiGenomics AG (DE)
JOURNAL
Location/Qualifiers
1..5491
location="synthetic construct"
/db_xref="taxon:32630"
/note="Chemically treated genomic DNA (Homo sapiens)"
BASE COUNT      1262 a      176 c      1573 g      2480 t
ORIGIN

Query Match      1.2%; Score 31; DB 6; Length 5491;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1952 ttgcgtcatccaaactaatcatcatcctca 1982
|||||
|||||
Db 1532 TTGCGTCATCTCCAACTAATCATTCA 1502

Search completed: June 11, 2002, 09:32:20
Job time: 9778 sec

```


C	10	330	13.1	554	22	AAK68349	Human immune/haema
C	11	260	10.3	1060	22	AAK57148	Human immune/haema
C	12	260	10.3	1090	22	AAK68654	Human immune/haema
C	13	260	10.3	1090	22	AAK68655	Human immune/haema
C	14	191	7.6	410	22	ABA08327	Human sensory GPCR
C	15	32	1.3	2771	21	AAZ50743	Rat sensory transd
C	16	31	1.2	5491	22	AAZ46390	Tumour suppressor
C	17	27	1.1	2579	21	AAZ50744	Mouse sensory tran
C	18	26	1.0	877	23	ABL20505	Drosophila melanog
C	19	26	1.0	1626	23	ABL23609	Drosophila melanog
C	20	26	1.0	2877	23	ABL20504	Drosophila melanog
C	21	26	1.0	3842	23	ABL23608	Drosophila melanog
C	22	26	1.0	5918	23	ABL12364	Drosophila melanog
C	23	25	1.0	1923	23	ABL22293	Drosophila melanog
C	24	25	1.0	3923	23	ABL22292	Drosophila melanog
C	25	24	1.0	201	23	AAK69964	DNA encoding novel
C	26	24	1.0	581	22	AAK92072	Human cDNA 5'-end
C	27	24	1.0	581	22	AAK93358	Human cDNA clone r
C	28	24	1.0	1251	19	AAV28701	Mouse apoptosis inh
C	29	24	1.0	1812	23	AAK91913	DNA encoding novel
C	30	24	1.0	1840	22	AAK94394	Human full-length
C	31	24	1.0	4260	23	ABL04783	Drosophila melanog
C	32	24	1.0	5651	22	ABA08791	Human d889M15.3 f
C	33	24	1.0	5653	22	AAH99780	Human protein encod
C	34	24	1.0	5653	22	AAI60854	Human polynucleot
C	35	24	1.0	5654	22	AAH99496	Human protein encod
C	36	24	1.0	5718	22	AAK57077	DNA encoding novel
C	37	24	1.0	5718	23	ABL04853	Drosophila melanog
C	38	24	1.0	7606	23	ABL04782	Drosophila melanog
C	39	24	1.0	9400	23	AAK57076	DNA encoding Drosoc
C	40	24	1.0	9400	23	ABL04852	Drosophila melanog
C	41	23	0.9	252	18	AAI65064	Canine genomic micr
C	42	23	0.9	252	20	AAK17801	Microsatellite rep
C	43	23	0.9	307	20	AAK232044	Human MEFH2 relat
C	44	23	0.9	307	22	AAK90301	HFKC19F cDNA clon
C	45	23	0.9	493	21	AAA37028	Human PRO831 (UNQ

PE 28-MAR-2001; 2001WO-US09996.
XX
PR 28-MAR-2000; 2000US-192978P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
PI Zambrowicz B, Sands AT.
XX
WPI: 2001-616474/71.
DR P-PSDB; AAE11969.

PR Novel isolated polynucleotides encoding human G protein coupled
 PN receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 PR expressed NGPCRs for diagnosis of disease, and as probes or primers -
 XX
 PS Claim 1; Page 71-72; 80pp; English.

CC The p187 protein sequence is a CDNA encoding human novel G-protein coupled
CC receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting
CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeats rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.

XX Sequence 2526 BP; 469 A; 770 C; 697 G; 587 T; 3 other;

Query Match	100.0%	Score 2526	DB 22	Length 2526
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2526; Conservative	0	Mismatches	0	Gaps 0

QY	1	atctgagctctgcaagagctgcgagctgtgctgagcttgagcttcaattctctgtctgagcc	60
Db	1	atcgagctctgcaagagctgcgagctgtgctgagcttgagcttcaattctctgtctgagcc	60
QY	61	tttgctctgcaatgacagagatctctccggaactcaacctcccggaagattacacctg	120
Db	61	tttgctctgcaatgacagagatctctccggaactcaacctcccggaagattacacctg	120
QY	121	gcaagagcgtgtccctctccatcttcctgagctgtgctgaggtgagagcaacaaccgaggtgacc	180
Db	121	gcaagagcgtgtccctctccatcttcctgagctgtgctgaggtgagagcaacaaccgaggtgacc	180
QY	181	ctgctgtacacaggtctctgtagcttcaatgtagcatgtagcatcaccccttccagagctatgacg	240
Db	181	ctgctgtacacaggtctctgtagcttcaatgtagcatgtagcatcaccccttccagagctatgacg	240
QY	241	cttgaggtctgagagataaacaactccacagccctgtgtgccaacatcacccctggagttac	300
Db	241	cttgaggtctgagagataaacaactccacagccctgtgtgccaacatcacccctggagttac	300
QY	301	cagctgtatgagtgtgttctgtaactctgcacaatgtgatatgccaagctggaagtgctctcc	360
Db	301	cagctgtatgagtgtgttctgtaactctgcacaatgtgatatgccaagctggaagtgctctcc	360
QY	361	ctgcagagggcaaacccacatagagcttccaaaggagaccttccactatccctctagcgtg	420
Db	361	ctgcagagggcaaacccacatagagcttccaaaggagaccttccactatccctctagcgtg	420
QY	421	ctggcagatgatttggcctgacagacccaacacctgtctgccaacacagcgcgcctgtgagc	480

Db	421	ctgacagatgattggccttgaaagacaaacggtgcgcgcacacacgcccctgctgtgac	480
Qy	481	cccttcctgtgtcccatgatgatactatgctgacagacgagacgctcagcgttgaacgag	540
Db	481	cccttcctgtgtcccatgatgatactatgctgacagacgagacgctcagcgttgaacgag	540
Qy	541	cagtatccctcttctctgctgcgacacatcccaatgaaatgaaagatgtagacatgtg	600
Db	541	cagtatccctcttctctgctgcgacacatcccaatgaaatgaaagatgtagacatgtg	600
Qy	601	ctgcgcgcgcgaagatctcgaggtgagaccctgatactctcctgtgttgacagcagttaacacat	660
Db	601	ctgcgcgcgcgaagatctcgaggtgagaccctgatactctcctgtgtgttgacagcagttaacacat	660
Qy	661	gggcagactaagggtgtgcagagcactgtgagaaaccagacacatgtgtacagggatctgtcatgtct	720
Db	661	gggcagactaagggtgtgcagagcactgtgagaaaccagacacatgtgtacagggatctgtcatgtct	720
Qy	721	ttcaagagacatcatgccccctctctctgcgcacaggtgagcagatgagagagatgagctcatgt	780
Db	721	ttcaagagacatcatgccccctctctctgcgcacaggtgagcagatgagagagatgagctcatgt	780
Qy	781	cgcacacctgcgcacacagcgcgcgggcccacccgctgtgtgttttttccagccgcgagcttgcc	840
Db	781	cgcacacctgcgcacacagcgcgcgggcccacccgctgtgtgttttttccagccgcgagcttgcc	840
Qy	841	agggtgttttttcgaatcccggtgtgctgtgacaaaccctgactgtgcagagttgtgtgtctccta	900
Db	841	agggtgttttttcgaatcccggtgtgctgtgacaaaccctgactgtgcagagttgtgtgtctccta	900
Qy	901	gaagcctgtgcgcctctccacagacacatcactggggtgtcccgagatccagcgcattgtgagt	960
Db	901	gaagcctgtgtccctctccacagacacatcactgtgggtgtcccgagatccagcgcattgtgagt	960
Qy	961	gtgcgcgcgcgtgtgcacatccagaaagaggtgtctccctgcgcctcaagagcgttttgaagaaagcc	1020
Db	961	gtgcgcgcgcgtgtgcacatccagaaagaggtgtctccctgcgcctcaagagcgttttgaagaaagcc	1020
Qy	1021	tatgcaccggtgcacaaagagagggcccttcagagcctgtgcacaaagagcctcctgtgtgcagacagc	1080
Db	1021	tatgcaccggtgcacaaagagagggcccttcagagcctgtgcacaaagagcctcctgtgtgcagacagc	1080
Qy	1081	aatcagctctctgcagaaatgacacagcttctatgrracacacagatgcccagaagcttcaaaagcc	1140
Db	1081	aatcagctctctgcagaaatgacacagcttctatgrracacacagatgcccagaagcttcaaaagcc	1140
Qy	1141	ttctccatgagtttctgctctcaacagcctaacgggcgtgtatgcggtgtgcccatagtgctc	1200
Db	1141	ttctccatgagtttctgctctcaacagcctaacgggcgtgtatgaggtgtgcccatagtgctc	1200
Qy	1201	caccagcctcctgcgcgtgtgtgctcctctgagagctgtgttccaaaggggcgcgttcaaccctgcgag	1260
Db	1201	caccagcctcctgcgcgtgtgtgctcctctgagagctgtgttccaaaggggcgcgttcaaccctgcgag	1260
Qy	1261	cttttgagacagatccacaaaggtgcatcttccctctctacaaagagacatgtgtgcgtttaat	1320
Db	1261	cttttgagacagatccacaaaggtgcatcttccctctctctacaaagagacatgtgtgcgtttaat	1320
Qy	1321	gacaaacagatccctccatgagctatataaataatgtcctgtgagatctgtgataccccaag	1380
Db	1321	gacaaacagatccctccatgagctatataaataatgtcctgtgagatctgtgataccccaag	1380
Qy	1381	tggagcttcaaggtgcctcgtgttccctccaatgtgtctccagttccagctaaacataatgag	1440
Db	1381	tggagcttcaaggtgcctcgtgttccctccaatgtgtctccagttccagctaaacataatgag	1440
Qy	1441	accaaaatccaggtgtgcacggaaagagaaacaaaggtgcttcaagtctgtgtgttccagcgac	1500
Db	1441	accaaaatccaggtgtgcacggaaagagaaacaaaggtgcttcaagtctgtgtgttccagcgac	1500
Qy	1501	tgtcttgaagggcacacagcagatgtgthaoggtttccatcactgtctgtttgaggtgtg	1560

```
Dh 1501 tgtctgaagggaaccagcgagtggtacgggttccatcactgctgcttgagtggtg 1560
Qy 1561 cccgttgagggttgagaccttctcaacaagtgacctctacagatgccaacctgtggg 1620
Db 1561 cccgttgagggttgagaccttctcaacaagtgacctctacagatgccaacctgtggg 1620
Qy 1621 aaagaaggttgagccaaggaagcagaacctgctcccgagcagctgtgtgtttg 1680
Db 1621 aaagaaggttgagccaaggaagcagaacctgctcccgagcagctgtgtgtttg 1680
Qy 1681 gcttgctgagcacaacctcttggtgctgcttggaagctaaacagctgctgctgctg 1740
Db 1681 gcttgctgagcacaacctcttggtgctgcttggaagctaaacagctgctgctgctg 1740
Qy 1741 ctgcttggaactgctgctgctgttgcctgagcactagacacctgtgtgtagtgacga 1800
Db 1741 ctgcttggaactgctgctgctgttgcctgagcactagacacctgtgtgtagtgacga 1800
Qy 1801 gggggccgctgctgcttctctatagtgggctccctgagcagagtgagtgagccttat 1860
Db 1801 gggggccgctgctgcttctctatagtgggctccctgagcagagtgagtgagccttat 1860
Qy 1861 ggtctcttgagggaaccaccaagaagcctgctgctgctgctgctgctgctgctt 1920
Db 1861 ggtctcttgagggaaccaccaagaagcctgctgctgctgctgctgctgctgctt 1920
Qy 1921 ggttccacctctctcctgctgctgctgctgctgctgctgctgctgctgctgctgct 1980
Db 1921 ggttccacctctctcctgctgctgctgctgctgctgctgctgctgctgctgctgct 1980
Qy 1981 aagtttccaccaaggtacacatcttcaacagcctggtgctcacaacacaggtgctgctg 2040
Db 1981 aagtttccaccaaggtacacatcttcaacagcctggtgctcacaacacaggtgctgctg 2040
Qy 2041 ctgttctgtagtgaactgaactgaagcggccagctgcttactgctgctgctgctgctg 2100
Db 2041 ctgttctgtagtgaactgaactgaagcggccagctgcttactgctgctgctgctgctg 2100
Qy 2101 tggagcccaactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2160
Db 2101 tggagcccaactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2160
Qy 2161 acaagaagcaactccctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2220
Db 2161 acaagaagcaactccctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2220
Qy 2221 agtgccttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2280
Db 2221 agtgccttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2280
Qy 2281 tgytcaacctcaagcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2340
Db 2281 tgytcaacctcaagcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2340
Qy 2341 agcgtctcaagcagcgaagtacactgctgctgctgctgctgctgctgctgctgctgctg 2400
Db 2341 agcgtctcaagcagcgaagtacactgctgctgctgctgctgctgctgctgctgctgctg 2400
Qy 2401 agcagcggtcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2460
Db 2401 agcagcggtcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2460
Qy 2461 ctcaacaagcagaagcacttccagcctcctcaagcactgagcagcagcagcagcagcagc 2520
Db 2461 ctcaacaagcagaagcacttccagcctcctcaagcactgagcagcagcagcagcagcagc 2520
Qy 2521 acctga 2526
Db 2521 acctga 2526
```

RESULT 2
AAD19505

```
ID AAD19505 standard; DNA: 2951 BP.
XX
AC AAD19505;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human novel G-protein coupled receptor DNA with 5' and 3' regions.
XX
KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW antisense-therapy; signal transduction; behavioural disorder; obesity;
KW heartbeat rate; inflammation; immune disorder; diabetes; cancer;
KW coronary disease; ds.
XX
OS Homo sapiens.
XX
PN WO200172842-A2.
XX
PD 04-OCT-2001.
XX
PE 28-MAR-2001; 2001WO-US09996.
XX
PR 28-MAR-2000; 2000US-192978P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
XX
DR WPI: 2001-616474/71.
XX
PT Novel isolated polynucleotides encoding human G protein coupled
PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
XX
PS Disclosure; Page 79-80; 80pp; English.
XX
CC The present sequence is human novel G-protein coupled receptor (NGPCR)
CC DNA with 5' and 3' regions. NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences, are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting
CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeat rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.
XX
SQ Sequence 2951 BP; 559 A; 888 C; 831 G; 669 T; 4 other;
```

Query Match 100.0%; Score 2526; DB 22; Length 2951;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 atgctgctgcaagcgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 60
Db 232 atgctgctgcaagcgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 291
Qy 61 ttgctgctgcaagcgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 120
Db 292 ttgctgctgcaagcgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 351
Qy 121 gaagcctgttccctccatctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 180
Db 352 gaagcctgttccctccatctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 411
```

QY 181 ctgtgtgacaggtctgttagcttcaatgagcaltggtctaccacctcttcaggtatgcg 240
|||||
Db 412 cgtgtgacaggtctgttagcttcaatgagcaltggtctaccacctcttcaggtatgcg 471
QY 241 ctgtgtgttgagagataaacaatccacgcgcctgtctcccaatcacctgggttac 300
|||||
Db 472 ctgtgtgttgagagataaacaatccacgcgcctgtctcccaatcacctgggttac 531
QY 301 cagctgtatgt 360
|||||
Db 532 cagctgtatgt 591
QY 361 ctgt 420
|||||
Db 592 ctgt 651
QY 421 ctgt 480
|||||
Db 652 ctgt 711
QY 481 ccttctctgt 540
|||||
Db 712 ccttctctgt 771
QY 541 cagtatccctcttctctgt 600
|||||
Db 772 cagtatccctcttctctgt 831
QY 601 ctgt 660
|||||
Db 832 ctgt 891
QY 661 ggt 720
|||||
Db 892 ggt 951
QY 721 ttcagaagacatcatgtccctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
|||||
Db 952 ttcagaagacatcatgtccctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1011
QY 781 cgtcacctgt 840
|||||
Db 1012 cgtcacctgt 1071
QY 841 aggt 900
|||||
Db 1072 aggt 1131
QY 901 gaagcctgt 960
|||||
Db 1132 gaagcctgt 1191
QY 961 ggt 1020
|||||
Db 1192 ggt 1251
QY 1021 tctgt 1080
|||||
Db 1252 tctgt 1311
QY 1081 aatcagctctgtcagaagaatgtcagaagcttctcagtcacacagatgtcccaagctca 1140
|||||
Db 1312 aatcagctctgtcagaagaatgtcagaagcttctcagtcacacagatgtcccaagctca 1371
QY 1141 tctctcagatgtctgtcctacacacacacacacacacacacacacacacacacacac 1200
|||||
Db 1372 tctctcagatgtctgtcctac 1431
QY 1201 caacagctctgt 1260
|||||
Db 1432 caacagctctgt 1491

QY 1261 ctttggagacagatccacaagaagtgtcatcttctctacacaagaacatgtgtgtttaat 1320
|||||
Db 1492 ctttggagacagatccacaagaagtgtcatcttctctacacaagaacatgtgtgtttaat 1551
QY 1321 gacaacagagatccctctcagtgatataacataatgtctgtgtgtgtgtgtgtgtgtgt 1380
|||||
Db 1552 gacaacagagatccctctcagtgatataacataatgtctgtgtgtgtgtgtgtgtgtgt 1611
QY 1381 tggacctcagcgttctctgt 1440
|||||
Db 1612 tggacctcagcgttctctgt 1671
QY 1441 accaaatccagatgtgcacggaagaagaacacacacacacacacacacacacacacac 1500
|||||
Db 1672 accaaatccagatgtgcacggaagaagaacacacacacacacacacacacacacacac 1731
QY 1501 tgtcttgaagggtcacaagagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560
|||||
Db 1732 tgtcttgaagggtcacaagagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1791
QY 1561 cctgt 1620
|||||
Db 1792 cctgt 1851
QY 1621 aaagaagaatgt 1680
|||||
Db 1852 aaagaagaatgt 1911
QY 1681 gcttgcgt 1740
|||||
Db 1912 gcttgcgt 1971
QY 1741 ctgt 1800
|||||
Db 1972 ctgt 2031
QY 1801 ggt 1860
|||||
Db 2032 ggt 2091
QY 1861 ggt 1920
|||||
Db 2092 ggt 2151
QY 1921 ggttccacatcttctctgt 1980
|||||
Db 2152 ggttccacatcttctctgt 2211
QY 1981 aagtttccacaagaagtacatcatcttccacacacacacacacacacacacacacac 2040
|||||
Db 2212 aagtttccacaagaagtacatcatcttccacacacacacacacacacacacacacac 2271
QY 2041 ctgt 2100
|||||
Db 2272 ctgt 2331
QY 2101 tggacccac 2160
|||||
Db 2332 tggacccac 2391
QY 2161 acaagac 2220
|||||
Db 2392 acaagac 2451
QY 2221 agtgccttgcctgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2280
|||||
Db 2452 agtgccttgcctgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2511
QY 2281 tgtgtac 2340
|||||
Db 2512 tgtgtac 2571
QY 2341 agcgttaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2400

Dh	781	gaagcctatgcccgggagagacaagggggcccttaaggcccttggcacaagaaggtccctgtgtgc	840
Oy	1075	agcagacaataagctcttgcagagaatgtccaaagcttccatgycacaagaatgcccagaatc	1134
Dh	841	agcagacaataagctcttgcagagaatgtccaaagcttccatgycacaagaatgcccagaatc	900
Oy	1135	aaagcctctcccaagaaattctgcctacaacgcataacggggctgtgtatgcygtggccat	1194
Dh	901	aaagcctctcccaagaaattctgcctacaacgcataacggggctgtgtatgcygtggccat	960
Oy	1195	ggcctccacaacagctccctcgggctgtgtccctctggagctgtgtccaaagggccagttacacc	1254
Dh	961	ggcctccacaacagctccctcgggctgtgtccctctggagctgtgtccaaagggccagttacacc	1020
Oy	1255	tggcagacttctggagcagatccacaagaatgcatlctcctctacaacaagaagacatgtgtcg	1314
Dh	1021	tggcagacttctggagcagatccacaagaatgcatlctcctctacaacaagaagacatgtgtcg	1080
Oy	1315	tttaatgacaacagagatcccccagtaagtataataataattgctctggagctgtgaatgga	1374
Dh	1081	tttaatgacaacagagatcccccagtaagtataataataattgctctggagctgtgaatgga	1140
Oy	1375	cccaagtgagacttccaaagcgtccctcggttccotcaaatggtctccaaagttcagctaaacata	1434
Dh	1141	cccaagtgagacttccaaagcgtccctcggttccotcaaatggtctccaaagttcagctaaacata	1200
Oy	1435	aatgagacccaanaalccagtggtgcacgggaagaagacaacagtgctctaagtctgtgtgtcc	1494
Dh	1201	aatgagacccaanaalccagtggtgcacgggaagaagacaacagtgctctaagtctgtgtgtcc	1260
Oy	1495	agcgacatgctcttgaagggcacaacggagatggtttaagggtttccatcactgctgtcttggag	1554
Dh	1261	agcgacatgctcttgaagggcacaacggagatggtttaagggtttccatcactgctgtcttggag	1320
Oy	1555	tgtctgccccttggggagcttggaagcttccctccacaagaagtgaaactataagatgtccagact	1614
Dh	1321	tgtgtgccccttggggagcttggaagcttccctccacaagaagtgaaactataagatgtccagact	1380
Oy	1615	tgtgaggaaagaagaatggtgcacctgtaggggaagcagaactgtctcccgcgcaatgtgtgtg	1674
Dh	1381	tgtgaggaaagaagaatggtgcacctgtaggggaagcagaactgtctcccgcgcaatgtgtgtg	1440
Oy	1675	tttttggccttgggtgagacaaactcttgggggtgtgtctgtgcagctctaaagctgtctgtg	1734
Dh	1441	tttttggccttgggtgagacaaactcttgggggtgtgtctgtgtgcagctctaaagctgtctgtg	1500
Oy	1735	ctgtgctgctgtgggaactgtgtgcgcgttcttgcctgagcaactagaaccccgtgtgtggag	1794
Dh	1501	ctgtgctgctgtgggaactgtgtgcgcgttcttgcctgagcaactagaaccccgtgtgtggag	1560
Oy	1795	tcaagcagggggccgcgtgtgtcttcttatatgtctgtgggtccctctggcagcaggtatgtgcagc	1854
Dh	1561	tcaagcagggggccgcgtgtgtcttcttatatgtctgtgggtccctctggcagcaggtatgtgcagc	1620
Oy	1855	ctctatgtgctcttctgggggaacccaagaagctgtgtgtgtctatgcgcgaagccctctt	1914
Dh	1621	ctctatgtgctcttctgggggaacccaagaagctgtgtgtgtctatgcgcgaagccctctt	1680
Oy	1915	ggccctgtgtttcacacatcttcctcgttcctgcctgcctgaagtttgcgtcatcttcaactaatc	1974
Dh	1681	ggccctgtgtttcacacatcttcctcgttcctgcctgcctgaagtttgcgtcatcttcaactaatc	1740
Oy	1975	atcttcaaggtttccacaagaagtaactacatlttccacaagcgtgtgtgtccaaaacacaggt	2034
Dh	1741	atcttcaaggtttccacaagaagtaactacatlttccacaagcgtgtgtgtgtccaaaacacaggt	1800
Oy	2035	gtctggcctgtttgtgtatgatacagctcaagcgcccaagctgtctatctgttcaactgtgtg	2094
Dh	1801	gtctggcctgtttgtgtatgatacagctcaagcgcccaagctgtctatctgttcaactgtgtg	1860
Oy	2095	gttgtgtgtgaccccaacagctgtgtcttaagggaataccagcgttccccaatctgtgtatgtt	2154
Dh	1861	gttgtgtgtgaccccaacagctgtgtcttaagggaataccagcgttccccaatctgtgtatgtt	1920

Oy	2215	gagtcgcagaagaccgaactccctcgggcttcatactgacctctctctacaaatgacctc	2214
Db	1921	gagtcgcagaagaccgaactccctcgggcttcatactgacctctctctacaaatgacctc	1980
Oy	2215	tcacatcagtcgcttctgcttcgaactacaccttggtaagagacttgcagagaaactaacagag	2274
Db	1981	tcacatcagtcgcttctgcttcgaactacaccttggtaagagacttgcagagaaactaacagag	2040
Oy	2275	gcacaatggttcaacttcagctcagctcgtccttcacaacttcgttctctgtatcgcttcttaac	2334
Db	2041	gcacaatggttcaacttcagctcagctcgtccttcacaacttcgttctctgtatcgcttcttaac	2100
Oy	2335	acggcagcagcttcaagcagcgaagtaacccgcttcgcgcccaaatgatgtgcggtgagc	2394
Db	2101	acggcagcagcttcaagcagcgaagtaacccgcttcgcgcccaaatgatgtgcggtgagc	2160
Oy	2395	agccttgagcagcgcttcggttggtaattcttcgtcctaagtgtacgtgatcccttcgcgc	2454
Db	2161	agccttgagcagcgcttcggttggtaattcttcgtcctaagtgtacgtgatcccttcgcgc	2220
Oy	2455	ccagacctcaacagcacagcagcacttcacggctccatctcagactaacagagcgctgc	2514
Db	2221	ccagacctcaacagcacagcagcacttcacggctccatctcagactaacagagcgctgc	2280
Oy	2515	ggctccacctga 2526	
Db	2281	ggctccacctga 2292	
RESULT 4			
ID	AAD17516		
AC	AAD17516	standard; cDNA; 2526 BP.	
XX	AAD17516;		
XX	10-DEC-2001	(first entry)	
DE	Human taste receptor, hTIR1 cDNA coding sequence.		
KW	Human; taste-cell-specific G protein-coupled receptor; hTIR1; drug;		
KW	genetic modulation; pharmaceutical; taste sensation; food industry;		
KW	chemosensory transduction; ss.		
XX	Homo sapiens.		
PH	Key	Location/Qualifiers	
FT	CDS	1..2526	
FT		/*tag= a	
FT		/product= "Human taste receptor, hTIR1 protein"	
FT		/transl_except= (pos:820..823, aa:Phe)	
FT		/note= "This codon has an apparent 1 nucleotide insertion	
FT		which alters the reading frame"	
XX	MO200166563-A2.		
XX	13-SEP-2001.		
XX			
PF	07-MAR-2001; 2001MO-US07265.		
XX			
PR	07-MAR-2000; 2000US-0187546.		
PR	07-APR-2000; 2000US-0195536.		
PR	06-JUN-2000; 2000US-0209840.		
PR	23-JUN-2000; 2000US-0214213.		
PR	17-AUG-2000; 2000US-0226448.		
PR	03-JAN-2001; 2001US-0259227.		
XX			
PA	(SENO-) SENOMYX INC.		
XX			
PI	Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;		
XX			
DR	WPI: 2001-582267/65.		
DR	P-PSDB; AAEL0372.		


```

Db 1740 gctgcttggaactgcgtgcgtgttgcctgcgcacacagacacccctgtgtgagcagc 1799
Oy 1800 aaggagcgccgctgtgtcttcttctatgtgtgtccctgcgcagcaggtagtgcagccttca 1859
Db 1800 aaggagcgccgctgtgtcttcttctatgtgtgtccctgcgcagcaggtagtgcagccttca 1859
Oy 1860 tggctctcttgaggaaacccaagaagcctgcgtgtctgtctacagcagccctcttgccct 1919
Db 1860 tggctctcttgaggaaacccaagaagcctgcgtgtctgtctacagcagccctcttgccct 1919
Oy 1920 tggcttaccatctctcctgcctgcctgcagaggtgcctatccaaataatcatctt 1979
Db 1920 tggcttaccatctctcctgcctgcctgcagaggtgcctatccaaataatcatctt 1979
Oy 1980 caagtttccacaaaggtacacatcttaccacagcctgtgtgttcaaacacaggtgtctg 2039
Db 1980 caagtttccacaaaggtacacatcttaccacagcctgtgtgttcaaacacaggtgtctg 2039
Oy 2040 cctgttctgcatgatcatcagctcagcgccagcctgtctatctgtctactgtgtgtgt 2099
Db 2040 cctgttctgcatgatcatcagctcagcgccagcctgtctatctgtctactgtgtgtgt 2099
Oy 2100 gtggaccaccatgcctgcctagaggaataacacagcctcccccacatctgtgtgtgtgt 2159
Db 2100 gtggaccaccatgcctgcctagaggaataacacagcctcccccacatctgtgtgtgtgt 2159
Oy 2160 caaagagacaaactcctcctgagcttactgtcctccttaccatgagcctcctccat 2219
Db 2160 caaagagacaaactcctcctgagcttactgtcctccttaccatgagcctcctccat 2219
Oy 2220 cagtgctcttgcctgcctagcctacacgtgtgaagacatgtccagagaaactacaagagccaa 2279
Db 2220 cagtgctcttgcctgcctagcctacacgtgtgaagacatgtccagagaaactacaagagccaa 2279
Oy 2280 atggtgtacactttagcctgtccttcaacttcgtgtcctgtgtgtgtgtgtgtgtgtgtgt 2339
Db 2280 atggtgtacactttagcctgtccttcaacttcgtgtcctgtgtgtgtgtgtgtgtgtgtgt 2339
Oy 2340 cagcgtctacagacgaagtaactcctgcctgcgcagcaacatgatgtgcgtgcagcct 2399
Db 2340 cagcgtctacagacgaagtaactcctgcctgcgcagcaacatgatgtgcgtgcagcct 2399
Oy 2400 gagcagcgtcctgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2459
Db 2400 gagcagcgtcctgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2459
Oy 2460 cctcaacagacagagacacttccagcctccatccagagcttaacagagcgtgcggtc 2519
Db 2460 cctcaacagacagagacacttccagcctccatccagagcttaacagagcgtgcggtc 2519
Oy 2520 cacttga 2526
Db 2520 cacttga 2526

```

```

RESULT 5
AAZ50745
ID AAZ50745 standard; cDNA; 2333 BP.
XX
AC AAZ50745;
XX
XX 31-MAY-2000 (first entry)
DT
XX
DE Human sensory transduction G-protein coupled receptor-B3 cDNA.
XX
XX Human; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
XX sensory cell; taste receptor cell; screen; taste modulator;
XX pharmaceutical; food industry; taste topographic map; tongue; ss.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH

```

```

FT CDS 1..2333
FT /tag= a
FT /product= "GPCR-B3"
FT /trans_except= (pos:865..866, aa:His)
PN WO200006592-A1.
XX
XX 10-FEB-2000.
XX
XX 27-JUL-1999; 99WO-US17099.
XX
XX 28-JUL-1998; 98US-0094465.
XX
XX (REGC ) UNIV CALIFORNIA.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
PI
XX WPI; 2000-205451/18.
XX DR P-PSDB; AAY45023.
XX
XX New isolated sensory transduction G-protein coupled receptor, useful
XX for developing products for use in studying and modulating the taste
XX transduction pathway.
XX
XX Claim 5; Pages 79-80; 83pp; English.
XX
XX The present sequence is a cDNA encoding a taste cell specific G-protein
XX coupled receptor, GPCR-B3 which is involved in sensory transduction.
XX This sequence was isolated from the human testis library.
XX GPCR-B3 is specifically expressed in foliate and fungiform cells, with
XX lower expression in circumvallate taste receptor cells of the tongue.
XX The present sequence is used to screen compounds that modulate sensory
XX signalling in taste cells, especially taste modulators useful in
XX pharmaceutical and food industries to customise taste. The sequence
XX can also be used as probe for identifying taste cells and
XX subsets of taste receptor cells such as foliate, fungiform and
XX circumvallate. Such probes are also useful to generate taste
XX topographic maps that elucidate the relationship between the taste
XX cells of the tongue and sensory neurons leading to taste centres
XX in the brain.
XX
XX Sequence 2333 BP; 445 A; 705 C; 646 G; 537 T; 0 other;
SQ

```

```

Query Match 41.5%; Score 1048; DB 21; Length 2333;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1808; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

```

```

Oy 709 atctgcatgtcttcaagaacatcatgcccctctcctgcaggtgtgtgtgtgtgtgtgtgtgt 768
Db 517 atctgcatgtcttcaagaacatcatgcccctctcctgcaggtgtgtgtgtgtgtgtgtgtgt 768
Oy 769 cagtgctctatgtgcacacctgcgcagcggcgccagcctgtgtgtgtgtgtgtgtgtgtgtgt 828
Db 577 cagtgctctatgtgcacacctgcgcagcggcgccagcctgtgtgtgtgtgtgtgtgtgtgtgt 828
Oy 829 cggcagcttgccagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 888
Db 637 cggcagcttgccagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 888
Oy 889 tgggtcgcctcagaagcctgtggcctctccagagcaacatcatgtgtgtgtgtgtgtgtgtgt 948
Db 697 tgggtcgcctcagaagcctgtggcctctccagagcaacatcatgtgtgtgtgtgtgtgtgtgt 948
Oy 949 cggatgtggatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1008
Db 757 cggatgtggatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1008
Oy 1009 ttggaagaagcctatcccgcgagcagaagaagggccctagagcctgtccacaagggctcc 1068
Db 817 ttggaagaagcctatcccgcgagcagaagaagggccctagagcctgtccacaagggctcc 1068

```

QY	1069	tgvgcagcagcaatcagctctcgagagaaatgccaagctttatgcaacagatgcc	1128
Db	876	tgvgcagcagcaatcagctctcgagagaaatgccaagctttatgcaacagatgcc	935
QY	1129	aagctcaagccttcctccatgagttctgccttaacaaagcaacagcgctgtagtcggtg	1188
Db	936	aagctcaagccttcctccatgagttctgccttaacaaagcaacagcgctgtagtcggtg	995
QY	1189	gcccattgctcccaacagcctccctgggctctgtgcctctgagact - tgtcccaaggcgaggt	1247
Db	966	gcccattgctcccaacagcctccctgggctctgtgcctct - gagctctgtcccaaggcgaggt	1054
QY	1248	ctacccctgagcagcttttttgagacagatccacaagatgcaattcctcttaacaaagacac	1307
Db	1055	ctacccctgagcagcttttttgagacagatccacaagatgcaattcctcttaacaaagacac	1114
QY	1308	tgtgagcttaatgacaacagagaatccctccagttagcttaacataatgtccctggagctg	1367
Db	1115	tgtgagcttaatgacaacagagaatccctccagttagcttaacataatgtccctggagctg	1174
QY	1368	gaatgagccccaagtgagacccctccacggttccttcggttcctccaaatgctccagttcaact	1427
Db	1175	gaatgagccccaagtgagacccctccacggttccttcggttcctccaaatgctccagttcaact	1234
QY	1428	aaacataaatgagacaaataactccagtgcagcggaagagacaaacagctgctaaagtctgt	1487
Db	1235	aaacataaatgagacaaataactccagtgcagcggaagagacaaacagctgctaaagtctgt	1294
QY	1488	gtgttcacagagactgttttgaaagggacacagcgagtggttaaggtttccatcaactctgtg	1547
Db	1295	gtgttcacagagactgttttgaaagggacacagcgagtggttaaggtttccatcaactctgtg	1354
QY	1548	ctttgagtgctgccccctgtgggctctggagaccttcctccaaagaaatgagcttaacagtg	1607
Db	1355	ctttgagtgctgccccctgtgggctctggagaccttcctccaaagaaatgagcttaacagtg	1414
QY	1608	ccagaccttgtaggaagaagaatgtggcaccttgagggaaagcagacacgtcttcocgagac	1667
Db	1415	ccagaccttgtaggaagaagaatgtggcaccttgagggaaagcagacacgtcttcocgagac	1474
QY	1668	tgvtggttttttgagcttttgctgtagacaacactcttggtgtgctgcctggagacgttaaaagct	1727
Db	1475	tgvtggttttttgagcttttgctgtagacaacactcttggtgtgctgcctggagacgttaaaagct	1534
QY	1728	gctctctctgtctgtctcttgaggactctcgagcctgtttgctcttgagacctaagaaacccctgt	1787
Db	1535	gctctctctgtctgtctcttgaggactctcgagcctgtttgctcttgagacctaagaaacccctgt	1594
QY	1788	ggtagagctcagcaggggggcgcgactgtgcttcttaatgctgaggtccctcgagcaggtag	1847
Db	1595	ggtagagctcagcaggggggcgcgactgtgcttcttaatgctgaggtccctcgagcaggtag	1654
QY	1848	tgagcagcctctatgagctcttttgagggaacccaagaagcctgcgtgttgctaagcgcaagc	1907
Db	1655	tgagcagcctctatgagctcttttgagggaacccaagaagcctgcgtgttgctaagcgcaagc	1714
QY	1908	cctcttgcccttggtttcaacacatctccctgcctccctcgagctgagcttgctaatccaact	1967
Db	1715	cctcttgcccttggtttcaacacatctccctgcctccctcgagctgagcttgctaatccaact	1774
QY	1968	aatcatcatctcaagtttctccacaagaagttaactaatctcaacacagcctgtgtccaaaa	2027
Db	1775	aatcatcatctcaagtttctccacaagaagttaactaatctcaacacagcctgtgtccaaaa	1834
QY	2028	ccaagctgtgcctgtctgtttgtatgatcaagctcaagcgcccaagctgtctatctgtctaac	2087
Db	1835	ccaagctgtgcctgtctgtttgtatgatcaagctcaagcgcccaagctgtctatctgtctaac	1894
QY	2088	ttggctctgtgtgtgagacccaatctgcctctgaggaataacagcgcttcaccaatctgtgt	2147
Db	1895	ttggctctgtgtgtgagacccaatctgcctctgaggaataacagcgcttcaccaatctgtgt	1954
QY	2148	gagctcttgagtgcaagagacaaactccctctgggcttcatatactgagccttccttaacaatg	2207

Db	1955	gacgtctgagcgccacagagacacactcccggtgctctcaactgagcctctctactaaatg	2014
QY	2208	cctccctccactaaagtgaccttctgctgcagctacacccggtgaaggacttgcagagaacta	2267
Db	2015	cctccctctctcaagtgaccttctgctgcagctacacccggtgaaggacttgcagagaacta	2074
QY	2268	caacgagcgccaatggtgtcatcctctcagccctgcctctcaactctgctctgatacgcct	2337
Db	2075	caacgagcgccaatggtgtcatcctctcagccctgcctctcaactctgctctgatacgcct	2134
QY	2338	cttcaacacggtccagcgcttctactacgagcgaatgaactctgcctgcgagcaaatgatgtcgtg	2387
Db	2135	cttcaacacggtccagcgcttctactacgagcgaatgaactcctgcctgcgagcaaatgatgtcgtg	2194
QY	2388	gctgtagcagccctgagcagcggtcttcggttggtattttctgcctaaagtgtactagtatcct	2447
Db	2195	gctgtagcagccctgagcagcggtcttcggttggtattttctgcctaaagtgtactagtatcct	2254
QY	2448	ctgcgcgcacagacctcaaacagcaacagcaactctcagagcctccatctcagagactaacagag	2507
Db	2255	ctgcgcgcacagacctctcaacagcaacagcaactctcagagcctccatctcagagactaacagag	2314
QY	2508	gcgcgtgcggtcccaactctga	2526
Db	2315	gcgcgtgcggtcccaactctga	2333

RESULT	6
AAD19503	ID AAD19503 standard; cDNA; 1101 BP.
AC	AAD19503;
DT	18-DEC-2001 (first entry)
DE	Human novel G-protein coupled receptor (NGPCR) cDNA #3.
KW	Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW	antisense-therapy; signal transduction; behavioural disorder; obesity;
KW	heartbeats rate; inflammation; immune disorder; diabetes; cancer;
KW	coronary disease; ss.
OS	Homo sapiens.
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1101
FT	/tag= a
FT	/product= "Human NGPCR protein"
FT	520..522
FT	/tag= b
FT	/note= "Encodes Ala"
XX	WO200172842-A2.
PD	04-OCT-2001.
XX	
PF	28-MAR-2001; 2001WO-US09996.
PR	28-MAR-2000; 2000US-192978P.
PA	(LEXI-) LEXICON GENETICS INC.
P1	Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
P1	Zambrowsicz B, Sands AT.
DR	WPI; 2001-616474/71.
DR	P-PSDB; AAEL1971.
PT	Novel isolated polynucleotides encoding human G protein coupled
PT	receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
XX	expressed NGPCRs for diagnosis of disease, and as probes or primers -

CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heart rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.

xx
xx Sequence 705 BP; 130 A; 231 C; 170 G; 173 T; 1 other;

Query Match 27.9%; Score 705; DB 22; Length 705;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1822 atgctggctccctggcagcagtagtgagcgcctctatgctctcttggggaaccaca 1881
|||
Db 1 atgctggctccctggcagcagtagtgagcgcctctatgctctcttggggaaccaca 60

Oy 1882 aggcctgctgtgtgtgtagcgcagcgcctcttggccttggttccacatctctgtcc 1941
|||
Db 61 aggcctgctgtgtgtgtagcgcagcgcctcttggccttggttccacatctctgtcc 120

Oy 1942 tgcctgacagttgcgtcatctccaatacatcatctcaagtttccacaagaagttact 2001
|||
Db 121 tgcctgacagttgcgtcatctccaatacatcatctcaagtttccacaagaagttact 180

Oy 2002 acattctacacagcctgggtccaaaacacaggtgtgctgctgttgtatgataagctca 2061
|||
Db 181 acattctacacagcctgggtccaaaacacaggtgtgctgctgttgtatgataagctca 240

Oy 2062 ggcggccagcgtgtatctgtcttaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2121
|||
Db 241 ggcggccagcgtgtatctgtcttaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300

Oy 2122 gaataccagcgttcccccacatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2181
|||
Db 301 gaataccagcgttcccccacatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360

Oy 2182 ttcaactcgtcctctctctcaaatgtgctctctccacatgtgcttggcttgagctac 2241
|||
Db 361 ttcaactcgtcctctctctcaaatgtgctctctccacatgtgcttggcttgagctac 420

Oy 2242 ctgggtgaagagcttgcagagagaactacaagagagcaaatgtgtactcttgaagctgtc 2301
|||
Db 421 ctgggtgaagagcttgcagagagaactacaagagagcaaatgtgtactcttgaagctgtc 480

Oy 2302 ttcaactcgtcctctctctcaaatgtgctctctccacatgtgcttgaagagcaagttac 2361
|||
Db 481 ttcaactcgtcctctctctcaaatgtgctctctccacatgtgcttgaagagcaagttac 540

Oy 2362 ctgctcgtgcgccaacatgt 2421
|||
Db 541 ctgctcgtgcgccaacatgt 600

Oy 2422 ttctcgtcctaagttgtatgtatctcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2481
|||
Db 601 ttctcgtcctaagttgtatgtatctcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 660

Oy 2482 caggcctccatcagactacagagagcgtctgagctcactctga 2526
|||
Db 661 caggcctccatcagactacagagagcgtctgagctcactctga 705

RESULT 8
AAD17515
ID AAD17515 standard; DNA; 8191 BP.

xx
AC AAD17515;
xx
DT 10-DEC-2001 (first entry)
xx
DE Human taste receptor, hT1R1 full-length genomic DNA.
xx
KW Human: taste-cell-specific G protein-coupled receptor; T1R1; drug;
KW genetic modulation; pharmaceutical; taste sensation; food industry;
KW chemosensory transduction; ds.
OS
XX Homo sapiens.
FH
FH Key
FH CDS
FT
FT 1..7771
FT /tag= a
FT /product= "Human taste receptor, hT1R1 protein"
FT 1..846
FT /tag= b
FT /number= 1
FT /cons_splice= (5'site:NO, 3'site:NO)
FT 847..1037
FT /tag= c
FT /number= 1
FT 1038..1593
FT /tag= d
FT /number= 2
FT 1594..1900
FT /tag= e
FT /number= 2
FT 1901..2819
FT /tag= f
FT /number= 3
FT 2820..3581
FT /tag= g
FT /number= 3
FT 3582..4603
FT /tag= h
FT /number= 4
FT 4604..4816
FT /tag= i
FT /number= 4
FT 4817..5138
FT /tag= j
FT /number= 5
FT 5139..5259
FT /tag= k
FT /number= 5
FT 5260..6841
FT /tag= l
FT /number= 6
FT 6842..7771
FT /tag= m
FT /number= 6
FT 7772..8191
FT /tag= n
FT /number= 7
FT /cons_splice= (5'site:NO, 3'site:NO)
PN MO200166563-A2.
PD 13-SEP-2001.
PE 07-MAR-2001; 2001WO-US07265.
XX
XX 07-MAR-2000; 2000US-0187546.
PR 07-APR-2000; 2000US-0195536.
PR 06-JUN-2000; 2000US-0209840.
PR 23-JUN-2000; 2000US-0214213.
PR 17-AUG-2000; 2000US-0226448.
PR 03-JAN-2001; 2001US-02359227.
PA (SENO-) SENOMYX INC.

endocannabinoids can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmune, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders.

SQ Sequence 1147 BP; 249 A; 322 C; 331 G; 245 T; 0 other;

Query Match	21.3%	Score 538	DB 22	Length 1147
Best Local Similarity	99.8%	Pred. No. 7.6e-247		
Matches 588	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

QY	1006	gacgtttgaagaagcctatgtcccggtggagagcaaggaagcccttaggccttgcacaagaaggc	1065
Db	89	gcgtttgaagaagcctatgtcccggtggagagcaaggaagcccttaggccttgcacaagaaggc	148
QY	1066	tcctctgtgcagcagcaatctcagctctctgcagaagatgcacaaagctttatgtrcacacagctg	1125
Db	149	tcctctgtgcagcagcaatctcagctctctgcagaagatgcacaaagctttatgtrcacacagctg	208
QY	1126	cccaagctcaaaagccttctccatgagttctgtccctacaaagcataccgggctgtgtatctg	1185
Db	209	cccaagctcaaaagccttctccatgagttctgtccctacaaagcataccgggctgtgtatctg	268
QY	1186	gtggcccatgtgcctcccaacacagctccctgggctctgtccctcgtgagagctgtgttcccaaggcgga	1245
Db	269	gtggcccatgtgcctcccaacacagctccctgggctctgtccctcgtgagagctgtgttcccaaggcgga	328
QY	1246	gtctacccctctggcagcgtttcttgtagcagatccacacaaggtgcatcttccctctacaacaagagac	1305
Db	329	gtctacccctctggcagcgtttcttgtagcagatccacacaaggtgcatcttccctctacaacaagagac	388
QY	1306	actgtgtgcgttttaatgacaaacagagatcccccctcagtagctataacataattgtccttggagac	1365
Db	389	actgtgtgcgttttaatgacaaacagagatcccccctcagtagctataacataattgtccttggagac	448
QY	1366	tggagatggaccacaagtgagaccttccacaaggtccctcggtttccctccacatgtgtccacgttttag	1425
Db	449	tggagatggaccacaagtgagaccttccacaaggtccctcggtttccctccacatgtgtccacgttttag	508
QY	1426	ctaaacaataatgagaccacaatctccagctgtgcacggcaaaagagacaacacagtgctccatagct	1485
Db	509	ctaaacaataatgagaccacaatctccagctgtgcacggcaaaagagacaacacagtgctccatagct	568
QY	1486	gtgtgtgtccagcagctgtccttgaagggcacacagcagatgtgttaagtggtttccatcatctgc	1545
Db	569	gtgtgtgtccagcagctgtccttgaagggcacacagcagatgtgttaagtggtttccatcatctgc	628
QY	1546	tgctcttgagtgctgtccctctgttggggtcttggagaccttccctcaacaagaatg	1594
Db	629	tgctcttgagtgctgtccctctgttggggtcttggagaccttccctcaacaagaatg	677

RESULT	10
AAK68439	
ID	AAK68439 standard; DNA; 554 BP.

AC AAK68439;

DT 06-NOV-2001 (first entry)

DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23251
XX	
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis; ds.

[illegible]

PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-02559678.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM.
XX
XX WPI; 2001-483426/52.
DR
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 23251; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 554 BP; 106 A; 168 C; 163 G; 117 T; 0 other;

Query Match 13.1%; Score 330; DB 22; Length 554;
Best Local Similarity 99.7%; Pred No. 2.1e-147;
Matches 380; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2146 gtagatgcttgatgacagagacacacactcctggtgcttatactatgcttcttacaat 2205
1 gtagatgcttgatgacagagacacacactcctggtgcttatactatgcttcttacaat 60
DB 61 ggcctctcttcacatcagtccttgcctgctgagctacacggtgtaagacttgcacaggaac 120
QY 2206 ggcctctcttcacatcagtccttgcctgctgagctacacggtgtaagacttgcacaggaac 2265
1 ggcctctcttcacatcagtccttgcctgctgagctacacggtgtaagacttgcacaggaac 120
DB 61 ggcctctcttcacatcagtccttgcctgctgagctacacggtgtaagacttgcacaggaac 120
QY 2266 tacacagagacacaaatggttccacactcagcgtccttcaacttcgtcctgagatgcc 2325
121 tacacagagacacaaatggttccacactcagcgtccttcaacttcgtcctgagatgcc 180
DB 121 tacacagagacacaaatggttccacactcagcgtccttcaacttcgtcctgagatgcc 180
QY 2326 ttcttcacacagcgcagcgttactagagcaagtaacttcgtcctgagatgcc 2385
181 ttcttcacacagcgcagcgttactagagcaagtaacttcgtcctgagatgcc 240
DB 181 ttcttcacacagcgcagcgttactagagcaagtaacttcgtcctgagatgcc 240
QY 2386 gggctgagacagcctgagcagcgttcgttggttatttctgcttaagctgacgtgac 2445
1 gggctgagacagcctgagcagcgttcgttggttatttctgcttaagctgacgtgac 2445
DB 241 gggctgagacagcctgagcagcgttcgttggttatttctgcttaagctgacgtgac 300
QY 2446 cctctgcgcacacactcacaacagacagacagcacttcagcgtcctcaatgagactacag 2505
1 cctctgcgcacacactcacaacagacagacagcacttcagcgtcctcaatgagactacag 2505
DB 301 cctctgcgcacacactcacaacagacagacagcacttcagcgtcctcaatgagactacag 360
QY 2506 aggcgtctgcgctcacctga 2526
1 aggcgtctgcgctcacctga 2526
DB 361 aggcgtctgcgctcacctga 381

RESULT 11
AAK57148/C
ID AAK57148 standard; cDNA; 1060 BP.
XX
AC AAK57148;
XX
DT 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen encoding cDNA seq ID NO:2208.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214866.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.

PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241121.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPL: 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX
PS Disclosure; SEQ ID NO 24466; 3071bp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 1090 BP; 257 A; 283 C; 284 G; 266 T; 0 other;

Query Match 10.3%; Score 260; DB 22; Length 1090;
Best Local Similarity 99.7%; Pred. No. 5.9e-114;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 acaggtcttgtagcttcataatgacatgctacacaccttccacagctatgagcttggg 247
Dy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 753 acaggtcttgtagcttcataatgacatgctacacaccttccacagctatgagcttggg 812
QY 248 ttgaggagataaacaactccacggccctgctgccaacatacaccctgggtacacagct 307
Dy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 813 ttgaggagataaacaactccacggccctgctgccaacatacaccctgggtacacagct 872
QY 308 atgatgtgttctgacttgcacatgtaatagtacacagctgagatgctctccctgcag 367
Dy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 873 atgatgtgttctgacttgcacatgtaatagtacacagctgagatgctctccctgcag 932
QY 368 ggcaacacacatagagctccaagagagaccttccactatccctcctaagtgctggag 427
Dy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 933 ggcaacacacatagagctccaagagagaccttccactatccctcctaagtgctggag 992
QY 428 tgattggcctgacagacacacacgctgctgacacacacgagccctgctgagcccttcc 487
Dy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 993 tgattggcctgacagacacacacgctgctgacacacacgagccctgctgagcccttcc 1052
QY 488 tgggtgcccatg 498
Dy ||||||||||||
Dy 1053 tgggtgcccatg 1063

RESULT 13
AAK69655 standard; DNA: 1090 BP.
ID AAK69655
XX

AC AAK69655;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:24467.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN W0200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229387.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-02331968.
PR 14-SEP-2000; 2000US-0233297.
PR 14-SEP-2000; 2000US-0233298.
PR 14-SEP-2000; 2000US-0233299.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233400.

PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0235837.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249247.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.

PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX Disclosure: SEQ ID NO 24467; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK4703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
CC represent sequences used in the exemplification of the present invention.
XX
XX Sequence 1090 BP; 258 A; 283 C; 283 G; 266 T; 0 other:

Query Match 10.3%; Score 260; DB 22; Length 1090;
Best Local Similarity 99.7%; Pred. No. 5.9e-114;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 188 acaggtcttgagctcaatgagcagtgctaccactcttcacgctatgctggtggg 247
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 753 acaggtcttgagctcaatgagcagtgctaccactcttcacgctatgctggtggg 812
OY 248 ttgagagataaacaacacacacgagccctgctgccaacatcccttggtgacagctgt 307
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 813 ttgagagataaacaacacacacgagccctgctgccaacatcccttggtgacagctgt 872
OY 308 atgagtgtgttctgactctgcacatgtgtatgacagcgtgagatgctctccctgcag 367
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 873 atgagtgtgttctgactctgcacatgtgtatgacagcgtgagatgctctccctgcag 932
OY 368 ggcacacacataagcttccaaagagagaccttccactatccctccctacggtgctggag 427
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 933 ggcacacacataagcttccaaagagagaccttccactatccctccctacggtgctggag 992
OY 428 ttattgggctctgacagcacaacgctgctgcacacacagccgctgctgagcccttcc 487
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 993 tgattgggctctgacagcacaacgctgctgcacacacagccgctgctgagcccttcc 1052
OY 488 tgggtgcccatg 498
DB ||||||||||||
DB 1053 tgggtgcccatg 1063

RESURF 14
ABA08327 standard; cDNA; 410 BP.
ID ABA08327

XX ABA08327;
AC 11-JAN-2002 (first entry)
XX
DT Human sensory GPCR B3 homologue-encoding cDNA, SEQ ID NO:103.
XX
DE Human sensory GPCR B3 homologue-encoding cDNA, SEQ ID NO:103.
XX
XX Human: cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
KW cyostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antulcer; ss.
XX
XX Homo sapiens.
OS
XX WO200157188-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US03800.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Dermanac RT;
PI WPI, 2001-457740/49.
XX
XX P-PSDB; ABB11083.
DR
XX
XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
XX
XX Claim 1; Page 377; 1963pp; English.
XX
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,

CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.
 CC
 SQ Sequence 410 BP; 74 A; 135 C; 99 G; 102 T; 0 other;
 XX
 Query Match 7.6%; Score 191; DB 22; Length 410;
 Best Local Similarity 99.6%; Pred. No. 5.8e-81;
 Matches 241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 247 gtgagagagtaaacactccacgagccctctgcccacatcacctgggtaccagctg 306
 DB 147 gtgagagagtaaacactccacgagccctctgcccacatcacctgggtaccagctg 206
 QY 307 tatgatctgtctgtgactctgccaatgtatgccacgctgagatgctctccctgcca 366
 DB 207 tatgatctgtctgtgactctgccaatgtatgccacgctgagatgctctccctgcca 266
 QY 367 gggcaacacacacatagagctcccaagagacctctccactatccctacggtgctgca 426
 DB 267 gggcaacacacacatagagctcccaagagacctctccactatccctacggtgctgca 326
 QY 427 gtatattggcgctgacgacacacacgctgctgccacacacgcccctgctgagcccttc 486
 DB 327 gtatattggcgctgacgacacacacgctgctgccacacacgcccctgctgagcccttc 386
 QY 487 ct 488
 DB 387 ct 388
 RESULT 15
 AA250743
 ID AA250743 standard; cDNA; 2771 BP.
 XX
 AA250743;
 AC 31-MAY-2000 (first entry)
 DT
 XX
 DE Rat sensory transduction G-protein coupled receptor-B3 cDNA.
 XX
 KW Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
 KW sensory cell; taste receptor cell; screen; taste modulator;
 KW pharmaceutical; food industry; taste topographic map; tongue; ss.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 57..2579
 FT /*tag= a
 FT /product= "GPCR-B3"
 XX
 PN MO200006592-A1.
 XX
 PD 10-FEB-2000.
 XX
 PF 27-JUL-1999; 99WO-US17099.
 XX
 PR 28-JUL-1998; 98US-0094465.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
 XX
 DR WPI; 2000-205451/18.

DR P-PSDB; AA145021.
 XX
 CC New isolated sensory transduction G-protein coupled receptor, useful
 PT for developing products for use in studying and modulating the taste
 PT transduction pathway -
 XX
 PS Claim 5; Pages 76-77; 83pp; English.
 XX
 CC The present sequence is a cDNA encoding a taste cell specific G-protein
 CC coupled receptor, GPCR-B3 which is involved in sensory transduction.
 CC This sequence was isolated from the 12AP rat circumvallate cDNA library.
 CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with
 CC lower expression in circumvallate taste receptor cells of the tongue.
 CC The present sequence is used to screen compounds that modulate sensory
 CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customise taste. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.
 SQ Sequence 2771 BP; 564 A; 797 C; 754 G; 656 T; 0 other;
 XX
 Query Match 1.3%; Score 32; DB 21; Length 2771;
 Best Local Similarity 100.0%; Pred. No. 5.6e-05;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2119 agggataaccagcgcttcccccactctgtgat 2150
 DB 2172 agggataaccagcgcttcccccactctgtgat 2203

Search completed: June 11, 2002, 09:19:48
 Job time: 5800 sec